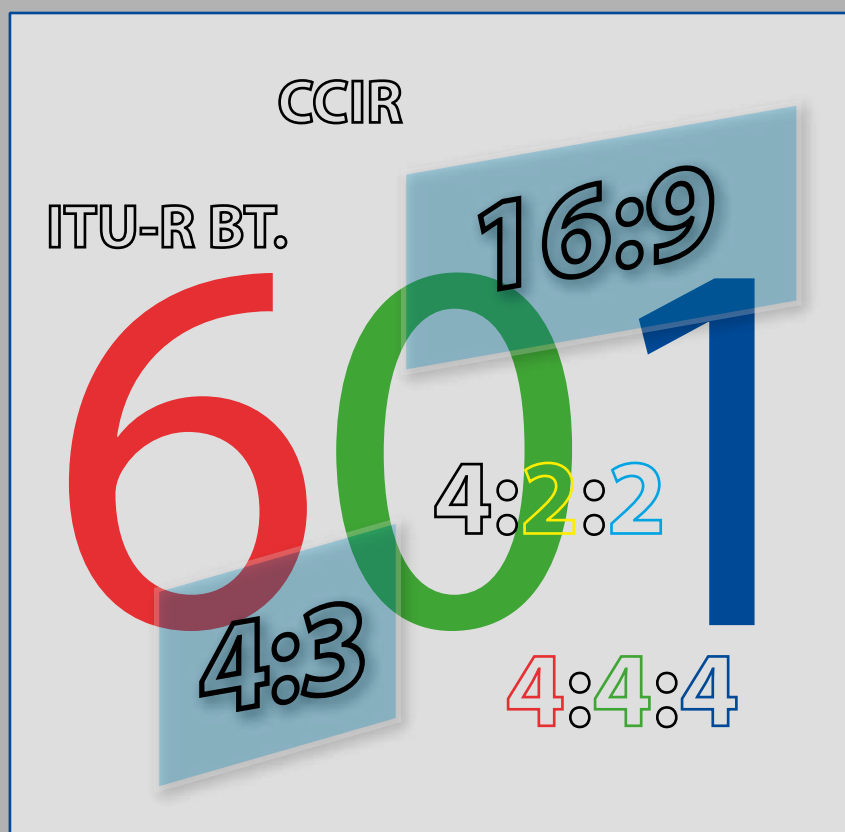
Sei Zero Uno...

e il segnale video TV diventa binario

Serie di articoli, pubblicati in più numeri di Elettronica e Telecomunicazioni, trattano e approfondiscono una singola tematica. Lo scopo dell'iniziativa **LeMiniSerie** è di raccogliere tali articoli, con una veste tipografica unitaria che ne faciliti la consultazione e apportando correzioni e aggiornamenti ritenuti opportuni.

Il secondo di questi volumi è intitolato Sei Zero Uno, il numero che identifica la norma internazionale alla base del passaggio dal segnale video analogico a quello numerico e raccoglie quattro articoli pubblicati nei numeri di aprile 1982, aprile 2003, dicembre 2006, dicembre 2009.

LeMiniSerie sono una iniziativa del
Centro Ricerche e Innovazione Tecnologica della Rai
www.crit.rai.it

In copertina:
Alcuni dei numeri e parametri che caratterizzano il passaggio dalla TV analogica a quella digitale.

Questo volume comprende 4 articoli pubblicati su Elettronica e Telecomunicazioni.

- *Gianfranco Barbieri:"Codifica Numerica del Segnale Video: standard per gli studi televisivi", n.2, 1982.*
- *Marzio Barbero e Natasha Shpuza: "Le origini del video digitale (La raccomandazione ITU-R BT.601)", aprile 2003*
- *Massimo Visca: "Formato d'immagine 16:9. Problemi di conversione", dicembre 2009*
- *Marzio Barbero e Natasha Shpuza: "Interfacce video (SDI, SDTI, ASI, HD-SDI, DVI, HDMI) ", dicembre 2006*

L'evoluzione dei sistemi di codifica digitale dell'informazione video ha origine con l'accordo internazionale, trenta anni fa, per convergere su un unico formato televisivo digitale: quello normalizzato nella attuale Raccomandazione ITU-R BT.601.

L'articolo del 1982, riprodotto come primo capitolo, rappresenta una completa e competente analisi della prima versione della nuova raccomandazione, ancora indicata provvisoriamente come CCIR AA/11.

Alla versione definitiva vene assegnato il numero 601; la denominazione attuale è del 1992, quando il CCIR divenne ITU-R e il prefisso BT indica che questa specifica si riferisce alla radiodiffusione (Broadcasting) televisiva (Television).

Il secondo capitolo è la versione aggiornata dell'articolo del 2003. Riassume brevemente i parametri principali della raccomandazione: essa fu definita con lo scopo di avere la maggior parte dei parametri in comune nel caso dei formati a 625 e 525 righe, con formato d'immagine 4:3, al fine di consentire economie di scala nella realizzazione degli apparati e di facilitare lo scambio internazionale dei programmi. Successivamente furono aggiunte le specifiche per i membri della famiglia destinati alla codifica del formato d'immagine 16:9, che caratterizzava i formati HDTV, ma che offriva innegabili vantaggi anche nel caso di definizione standard. Tale aggiunta fu però abbandonata, perché praticamente inutilizzata in produzione, infatti i sistemi di codifica MPEG, destinati alla diffusione e distribuzione (DVB e DVD), hanno preferito adottare come parametri di base del campionamento dell'immagine quelli previsti dalla Rac. 601 per il formato d'immagine 4:3, anche nel caso di formato 16:9.

E' proprio il formato d'immagine 16:9 ad essere oggi sempre più diffuso, e la coesistenza con il formato 4:3 è causa di problemi di conversione del rapporto di forma: l'ampia casistica di situazioni è oggetto del terzo capitolo, articolo pubblicato nel 2009.

Il rapporto di forma 4:3 è in uso fin dagli anni 50, dalla nascita del sistema televisivo, e ancora oggi viene normalmente utilizzato per la produzione e la trasmissione su tutte le reti Rai, costituisce inoltre la quasi totalità del materiale di archivio. Fino a pochi anni or sono, il parco ricevitori domestici era quasi interamente costituito da televisori in formato 4:3 e tale fatto ha impedito l'avvento del formato 16:9: anche la produzione in formato panoramico è sempre stata molto limitata. Questa situazione è cambiata con la penetrazione sul mercato dei display a schermo piatto, in genere in formato 16:9, e con l'introduzione della piattaforma di trasmissione digitale terrestre. La coesistenza di segnali e display con rapporti di forma diversi genera un'ampia casistica di situazioni in cui occorre convertire un segnale dal rapporto di forma 4:3 al 16:9, o viceversa. Sono descritte nel terzo capitolo le varie modalità con cui vengono presentati i formati 4:3 e 16:9, si analizza in sintesi il concetto di ripresa protetta e si riassumono i criteri per il rispetto delle Safe Areas.

Infine, il quarto capitolo, aggiornamento di un articolo del 2006, parte dalla Raccomandazione ITU-R BT.656 relative alla interfaccia per consentire l'interconnessione degli apparati basati sulla ITU-R BT.601 per poi descrivere brevemente le interfacce utilizzate per interconnettere gli apparati video sia nell'ambito professionale che in quello consumer.

Torino, novembre 2010

Marzio Barbero e Natasha Shpuza

Indice

Acronimi e sigle

ACE	Automatic Content Enhancement
ASI	Asynchronous Serial Interface
ATM	Asynchronous Transfer Mode
CCIR	Comité Consultatif International pour la Radio (dal 1992 è denominata ITU-R)
CCITT	Comité Consultatif International Télé-phonique et Télégraphique (dal 1992 è denominata ITU-T)
CIE	Commission Internationale de l'Éclairage (http://www.cie.co.at)
CRC	Cyclic Redundancy Check
DCP	Digital Content Protection (www.digital-cp.com)
DRM	Digital Rights Management
DVB	Digital Video Broadcasting (www.dvb.org)
DVD	Digital Versatile Disc
DVI	Digital Visual Interface
EAV	End of Active Video
EICTA	European Industry Association for Infor-mation Systems (dal marzo 2009 rinominata DIGITALEU-ROPE, www.digitaleurope.org/)
ENG	Electronic News Gathering
EBU	European Broadcasting Union (www.ebu.ch)

FIR	Finite Impulse Response
HEC	HDMI Ethernet Channel
HDCP	High-bandwidth Digital Content Protec-tion (v. DCP)
ITU-R	International Telecommunication Union (www.itu.int)
NRZI	Non-Return to Zero Inver*t*
NTSC	National Television System(s) Committee
PAL	Phase Alternating Line
RGB	Red Green Blue
SAV	Start of Active Video
SCART	Syndicat des Constructeurs d'Appareils Radiorécepteurs et Téléviseurs
SECAM	Serial Data Transport Interface
SDI	Serial Data Interface
SDTI	Serial Data Transport Interface
SMPTE	Society of Motion Picture and Television Engineers (www.smpte.org)
UER	Unione Européeenne de Radio-Télèvi-sion (v. EBU, www.ebu.ch)

CODIFICA NUMERICA DEL SEGNALE VIDEO:
STANDARD PER GLI STUDI TELEVISIVI

G. F. BARBIERI (*)

SOMMARIO — Nel corso degli ultimi anni si è assistito ad una penetrazione sempre più intensiva, negli studi televisivi, di apparati basati sull'impiego delle tecniche numeriche. Onde evitare una incontrollata proliferazione di standard e per gettare le basi di un sistema televisivo nuovo, capace di fornire prestazioni adeguate alla evoluzione tecnologica in atto nella componentistica numerica, è stato prodotto, da parte di vari organismi, tra cui l'UER, un grosso sforzo in termini di ricerca e sperimentazione. Come risultato si è pervenuti alla recente approvazione da parte del CCIR della Raccomandazione AA/11 sul sistema di codifica numerica per gli studi. Il sistema proposto riguarda esclusivamente gli impianti di generazione dei programmi, mentre, per quanto concerne la trasmissione dell'utente, si prevede che per molti anni ancora essa sarà basata sulle attuali tecniche di tipo analogico. Vengono qui illustrate le specifiche del nuovo standard e i criteri che sono stati seguiti per la loro definizione.

*SUMMARY — **Specification of the CCIR digital coding standard for television studios.** In the last years, a more and more intensive penetration of digital equipment in the studio complex has been observed. To avoid uncontrolled proliferation of standards and to pave the way for a new TV system, appropriate to the technological evolution in the digital components, an important endeavour has been carried out by various Organisations, in particular by the E.B.U., which has turned out in the recent approval of CCIR Rec. AA/11 on digital coding for television studios. This coding system only applies to the production installations; in fact, as far as broadcasting is concerned, the utilization od the current systems, based on the analogue techniques, is still foreseen for many years. Finally this article deals with specifications of the new standard and criteria for its definition.*

1. Introduzione.

Nel corso della XV Assemblea Plenaria del CCIR, tenutasi a Ginevra nel febbraio '82, è stato approvato il testo della nuova Raccomandazione AA/11 sulle specifiche di base del sistema di codifica numerica per gli studi televisivi (bibl. 1). L'evento è destinato ad avere grande risonanza poiché, oltre ad evitare che nei futuri impianti numerici si verifichi una incontrollata proliferazione di standard (¹), apre le porte verso l'unificazione, a livello mondiale, dei sistemi televisivi; infatti, i parametri sui quali è stata conseguita l'unanimità dei consensi, sono adattabili ad entrambe le norme di scansione che tuttora dividono il mondo in aree operanti a 625 righe/50 Hz e 525 righe/60 Hz.

Fin dall'inizio degli anni '70 gli organismi televisivi e le industrie operanti nel settore degli apparati da ripresa professionali hanno mostrato un crescente interesse verso i possibili vantaggi offerti dall'impiego delle tecniche numeriche nella produzione dei programmi televisivi. In particolare, in Europa l'UER (Unione Europeénne de Radiodiffusion - Europa occidentale) ha svolto, tramite la propria Commissione Tecnica, una intensa azione di coordinamento degli studi effettuati nei vari laboratori degli enti televisivi ad essa associati (bibl. 2); tale azione è stata coronata da successo nel marzo 1981, con l'accordo raggiunto tra UER e SMPTE (Society of Motion Picture and Television Engineers) accordo a cui si sono subito associate altre importanti istituzioni tra cui l'OIRT (Organisation Internationale de Radiodiffusion et Télévision - Europa orientale) circa l'adozione di una frequenza di campionamento del segnale di luminanza pari a 13.5 MHz come base per fissare le specifiche di un sistema che fosse accettabile ad entrambi gli organismi.

Ripercorrendo le varie tappe che hanno portato alla definizione del nuovo standard numerico, si constata facilmente come le previste finalità d'impiego della televisione numerica abbiano subito una sostanziale evoluzione nel corso degli anni: intese dapprima come un mezzo per facilitare la trasparenza delle reti di collegamento e quindi per favorire lo scambio internazionale dei programmi, le tecniche numeriche sono state, progressivamente, considerate soprattutto un formidabile strumento per generare programmi con linguaggi artistico-espressivi nuovi e con un incremento qualitativo considerevole.

(*) Dott. Ing. Gianfranco Barbieri del Centro Ricerche RAI di Torino.

Dattiloscritto pervenuto il 5 marzo 1982.

(¹) Poiché il segnale video codificato in forma numerica, si presta ad essere sottoposto ad una certa varietà di trattamenti che non sarebbero praticamente possibili con il segnale nella versione analogica tradizionale, si sta assistendo, da qualche anno, ad una immissione sempre più frequente sul mercato di apparati analogici che, al loro interno, fanno ricorso ad una conversione analogico-numerica e viceversa. In altre parole, le aree numeriche esistenti all'interno di uno studio assumono già l'aspetto di «isole» il cui numero è destinato via via ad aumentare fino alla completa trasformazione dell'impianto in versione numerica. In assenza di una normativa, sarebbero inevitabilmente adottate, da parte dei costruttori, le specifiche di co-decodifica ritenute caso per caso più convenienti, con grave pregiudizio per le future possibilità di interconnessione diretta tra «isole» numeriche.

Il vertiginoso sviluppo delle tecnologie integrate verificatosi negli ultimi anni offre, infatti, possibilità sempre nuove al trattamento elettronico delle immagini, favorendo l'utilizzazione del mezzo televisivo non soltanto come veicolo di ripresa e trasmissione della realtà, ma anche come strumento di esplorazione e composizione di immagini originali.

I punti fondamentali che caratterizzano il nuovo sistema numerico sono sostanzialmente:

— la codifica del segnale televisivo nella forma delle tre componenti, anziché del segnale composito come è d'uso negli attuali impianti analogici;

— l'articolazione del sistema secondo una famiglia di standard compatibili, ciascuno dei quali corrispondente ad un determinato livello qualitativo che è funzione delle possibilità operative dell'impianto preposto al suo trattamento.

Tra i possibili membri della famiglia, solo uno, e precisamente quello individuato come lo standard principale da studio, è stato per il momento definito nei suoi parametri più importanti; circa gli altri standard della gerarchia si è unicamente affermato il principio della estensibilità della famiglia sia verso livelli qualitativi inferiori (qualora particolari applicazioni impongano un risparmio di capacità trasmissiva a scapito della qualità) sia verso livelli superiori (per eventuali produzioni di carattere complesso).

La conseguenza più immediata dell'approvazione della Raccomandazione AA/11 sarà indubbiamente rappresentata da una accelerazione del processo di introduzione delle tecniche numeriche negli studi televisivi; infatti, l'industria moltiplicherà i suoi sforzi per realizzare nuovi apparati da immettere sul mercato, essendo ora disponibile una normativa di base a cui fare riferimento e su cui concentrare l'attività di sviluppo. È opportuno, tuttavia, sottolineare che le specifiche finora approvate riguardano i parametri principali del sistema; sono tuttavia in corso studi, particolarmente in campo UER, aventi come obiettivo la definizione di altri importanti elementi quali il formato del registratore videomagnetico numerico e l'interfaccia di entrata/uscita degli apparati.

2. Esigenze imposte dalle moderne tecniche di produzione dei programmi.

Fino ad oggi, il sistema televisivo basato sulle tecniche analogiche ha offerto una gamma più o meno continua di livelli qualitativi dell'immagine in relazione alle differenti possibilità operative degli apparati disponibili nell'impianto (bibl. 3).

Il caso più significativo è rappresentato dalle riprese di eventi di attualità, comunemente denominate ENG ([2]), per le quali è condizione primaria l'impiego di apparati leggeri, compatti e robusti anche se a prezzo di una certa penalizzazione della qualità globale; per quanto concerne invece i programmi che rientrano nella normale produzione da «studio», essi vengono solitamente prodotti trattando il segnale secondo le stesse norme dello standard di trasmissione, cioè codificato PAL, SECAM o NTSC.

Le limitazioni tecniche degli standards di codifica a colori, in particolare le larghezze di banda relativamente modeste delle componenti di crominanza e la impossibilità pratica di separare, senza residue interferenze mutue, luminanza e crominanza, consentono di effettuare sul segnale soltanto un numero limitato di operazioni; pertanto, in taluni studi il segnale viene trattato nel formato «in componenti» cioè mantenendo separate le tre componenti primarie. L'elevato margine di qualità di quest'ultimo formato permette di realizzare una certa gamma di effetti speciali tra cui l'intarsio a chiave cromatica ([3]), tuttavia la complessità dell'impianto e l'impossibilità pratica di registrare il segnale su nastro videomagnetico ([4]), ne limitano l'impiego alle prime fasi del processo di generazione.

I vari livelli qualitativi normalmente coesistono all'interno dell'impianto, in quanto è prassi corrente trasferire il segnale da una categoria di apparati all'altra; ad esempio, un segnale a livello di composito, prima di essere inviato al registratore videomagnetico, può essere intarsiato con un altro segnale a livello ENG, proveniente da contributi esterni, utilizzando una chiave d'intarsio ricavata in un'area di apparati in cui il segnale è trattato a livello di componenti separate.

La prospettiva di una conversione tecnologica degli impianti basata sulla rappresentazione in forma numerica del segnale d'immagine offre l'occasione per reimpostare la filosofia d'impiego dei mezzi di produzione televisiva in modo che vengano superate alcune tra le più gravose limitazioni ereditate dal passato; ad esempio, risulta particolarmente attraente la soluzione di codificare il segnale direttamente in componenti (Y, R-Y, B-Y) e di mantenere di tale formato in tutto l'impianto eliminando i difetti derivanti dall'uso del composito.

L'opportunità di avere disponibili diversi livelli di qualità corrispondenti a differenti campi di impiego e quindi a differenti possibilità operative degli apparati esclude, di fatto, che un'unica norma di codifica numerica sia, di per sé, adeguata a soddisfare una gamma estesa di esigenze talvolta contrastanti; poiché è indispensabile che gli impianti numerici mantengano almeno lo stesso grado di flessibilità offerto dalle attuali installazioni analogiche, ne consegue che il nuovo sistema deve essere articolato secondo un insieme di standard mutualmente compatibili. In altre parole, deve risultare possibile realizzare un impianto numerico scegliendo gli apparati in modo che qualità e costo siano adeguati al campo di applicazione previsto; inoltre, i segnali debbono essere convertibili da un livello all'altro mediante interpolazioni semplici e senza che l'operazione comporti interferenze intrinseche al processo di trascodifica (è superfluo osservare che la perdita di qualità derivante dal passaggio ad un livello inferiore della famiglia è irreversibile e non può essere compensata con il ritorno al livello superiore).

Con riferimento allo sviluppo in atto di nuove tecniche di produzione, si sta accentuando la tendenza a trasferire il maggior numero possibile di operazioni di trattamento creativo dell'immagine dall'area della ripresa a quella cosiddetta di «postpro-

---

([2]) ENG è acronimo di Electronic News Gathering.

([3]) Comunemente denominato Chromakey. Tecnicamente è possibile ricavare la chiave del chromakey anche dal segnale a colori composito ma la qualità dell'intarsio risulta in tal caso alquanto inferiore.

([4]) Come è noto i registratori videomagnetici professionali, oggi disponibili sul mercato, accettano unicamente il segnale nel formato composito.

duzione » (bibl. 4, 5); gli obiettivi che si intende perseguire consistono sostanzialmente nella ottimizzazione dei costi globali di produzione e nella realizzazione di una maggior flessibilità artistica.

Per chiarire meglio il significato della parola « postproduzione », è utile fare riferimento alla prassi, fino ad oggi seguita, in base alla quale la generazione del programma avviene sostanzialmente durante la ripresa (in studio o in esterna) delle singole sequenze e la registrazione delle stesse su nastro videomagnetico; ulteriori operazioni differite nel tempo sono limitate al montaggio dei vari spezzoni (operazione divenuta oggi alquanto flessibile grazie all'uso del computer come organo di gestione del processo) o tutt'al più alla correzione del bilanciamento colorimetrico; sono da escludersi, dati i limiti dei sistemi analogici, interventi sul segnale tendenti a correggere o modificare in qualche modo l'aspetto artistico-espressivo delle immagini registrate.

Sfruttando le potenzialità degli apparati basati sulle tecniche numeriche è invece possibile sviluppare nuovi metodi di produzione dei programmi tendenti a ridurre i tempi di impegno dello studio su cui gravano i maggiori costi o delle attrezzature per riprese in esterna, limitando tali tempi allo stretto indispensabile per la recitazione degli attori, e rinviando ulteriori operazioni di trattamento creativo a fasi di lavorazione successive da effettuarsi, a posteriori (e cioè in « postproduzione »), sul segnale registrato.

Tra le operazioni che si prevede di poter attuare, in postproduzione, entro tempi brevi acquistano una certa rilevanza, il ckromakey e una gamma praticamente illimitata di effetti speciali, la grafica-elettronica (ad es. vari tipi di animazione), la ricomposizione dell'immagine, la colorazione artificiale di scene reali per l'ottenimento di particolari effetti artistici (bibl. 6).

In un futuro meno immediato, ma realisticamente non molto remoto, si prevede, inoltre di poter semplificare i problemi di scenografia, che contribuiscono ad appesantire i costi di produzione, utilizzando in taluni casi la ripresa di modellini in scala a cui sovrapporre, ad esempio tramite intarsio elettronico, le immagini degli attori o di altri elementi che contribuiscono a completare il programma; ciò comporterà ovviamente, problemi di prospettiva e, di conseguenza, richiederà sofisticati processi di alterazione della geometria delle immagini.

In ogni caso, gran parte delle sopracitate operazioni sono accompagnate inevitabilmente da una maggior evidenziazione dei difetti originari del segnale e pertanto da una percettibile diminuzione della qualità soggettiva valutata sul prodotto finale cioè sull'immagine pronta per essere trasmessa all'utente; è significativo il caso dello « zoom » elettronico in cui il processo di espansione amplifica e rende maggiormente visibili i difetti di risoluzione dell'immagine originaria.

Si osservi, a questo proposito, che, mentre il trattamento del segnale analogico attraverso una catena di apparati è inevitabilmente accompagnato da una degradazione cumulativa e la qualità finale dipende dal numero di operazioni e dalla complessità delle stesse (è tipico il caso dei riversamenti da registratore video magnetico), nel caso del segnale numerico non vi è relazione tra l'entità del trattamento subito e qualità finale potendosi ottimizzare i parametri del sistema di codifica in modo che le uniche limitazioni qualitative siano quelle presenti all'atto dell'organizzazione del messaggio alla sorgente.

In definitiva, occorre che i parametri del sistema siano definiti tenendo conto dei necessari margini di qualità affinché quest'ultima venga mantenuta a livelli adeguati fino al termine dell'intero processo di lavorazione.

3. Criteri per la scelta dei parametri del sistema.

Le ripercussioni derivanti dall'adozione, all'interno degli studi televisivi, di un determinato sistema di codifica numerica non rappresentano soltanto un fatto interno al processo di produzione del programma ma creano un sostanziale impatto su tutta la catena televisiva (bibl. 7); nello sforzo di pervenire alla definizione di uno standard efficace e, soprattutto, adeguato ai prevedibili sviluppi tecnologici dei prossimi anni, è stato necessario prendere in considerazione un gran numero di fattori suscettibili di influenzare sia la qualità, sia la possibilità pratica di sviluppare apparati competitivi, anche in termini economici, con quelli analogici finora impiegati (bibl. 8, 9).

3.1. Forma del segnale codificato

Sono stati proposti due differenti approcci al problema:

— *codifica per componenti:* la componente di luminanza (Y) e le due componenti differenza di colore (R-Y e B-Y) sono codificate separatamente e trasmesse come flussi binari multiplati in TDM (Time Division Multiplex);

— *codifica del segnale composito:* il segnale viene codificato direttamente nella sua forma composita (PAL, SECAM o NTSC) ed è trasmesso come unico flusso binario.

Il primo sistema presenta indubbi vantaggi rispetto al secondo, a parte i problemi inerenti il periodo di transizione dall'analogico al numerico, durante il quale, dato il persistere di residue isole analogiche nella catena, si dovrà inevitabilmente accettare un certo numero di conversioni componenti-composito e viceversa, in cascata. Le ragioni più sostanziali per una scelta a favore del primo sono (bibl. 10, 11):

— i recenti sviluppi tecnologici, particolarmente nel campo della integrazione su larga scala e della registrazione magnetica, hanno creato nuove e attrattive possibilità nel trattamento elettronico delle immagini; il segnale in forma composita rappresenterebbe in ogni caso una complicazione e un freno alle possibilità operative degli apparati numerici;

— la codifica per componenti crea, di fatto, una unificazione del sistema televisivo a livello mondiale, in quanto offre la possibilità di definire uno standard numerico adattabile ad entrambe le norme di scansione tuttora persistenti ([5]). In pratica sarà possibile progettare apparati aventi il massimo di circuiteria comune, situazione non realizzabile con gli standard analogici o con un eventuale standard

([5]) Come si vedrà più avanti, la frequenza di campionamento è stata scelta anche in base a criteri di compatibilità tra norme di scansione a 625 righe e norme a 525 righe.

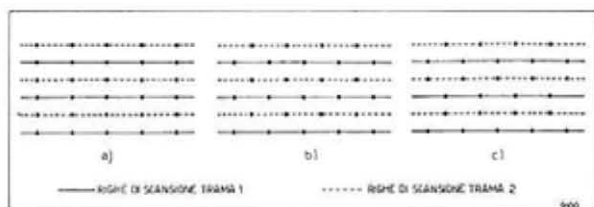



Fig. 1. — Le tre fondamentali strutture di campionamento. *a*) Ortogonale; *b*) Quincunx di trama; *c*) Quincunx di riga.

numerico in composito. In altre parole fatta eccezione per la frequenza di quadro (eventualmente commutabile), gli studi televisivi useranno in tutto il mondo lo stesso tipo di apparati con notevoli vantaggi per la produzione e lo scambio di programmi.

3.2. Struttura di campionamento e problemi di limitazione di banda

Come è noto, il processo di campionamento determina la formazione di spettri di modulazione centrati attorno ai multipli della frequenza di campionamento; condizione necessaria per il corretto recupero del segnale originario in banda base è che la frequenza di campionamento valga almeno il doppio del limite superiore della banda base (criterio di Nyquist) e pertanto le caratteristiche spettrali del segnale debbono venire definite con precisione prima che il campionamento abbia luogo.

Nel caso del segnale televisivo, derivante cioè da un processo di analisi sequenziale dell'immagine, si verifica usualmente una buona tenuta sotto controllo della larghezza di banda nella direzione orizzontale, mentre nella direzione verticale l'immagine viene, di fatto, campionata dalla scansione per righe prima che la distribuzione spettrale sia delimitata in modo adeguato; oltre a ciò, la susseguente riproduzione su cinescopio dell'informazione verticale risulta ulteriormente complicata dagli effetti derivanti dall'interallacciamento delle trame.

Tenuto conto dei vari problemi connessi con la struttura bidimensionale dell'informazione televisiva, si sono studiate varie strutture di campionamento nel tentativo di trovare la soluzione che garantisca la copertura più uniforme possibile dell'immagine da parte dei singoli campioni (bibl. 12). Nei casi in cui sussistano severe limitazioni al « bit rate » (ad esempio in applicazioni tipo ENG) sono risultate convenienti strutture di campionamento tipo « Quincunx » [6], come quelle illustrate in figura 1 *b* e *c*; la convenienza, deriva dal fatto che mediante opportuno filtraggio bidimensionale, la limitazione di banda (richiesta dal rispetto del criterio di Nyquist) interviene principalmente, sulle frequenze spaziali diagonali dell'immagine, mentre viene risparmiata parte della risoluzione orizzontale e verticale (bibl. 13).

Nel caso delle frequenze di campionamento proposte per lo standard di codifica da studio (a cui corrispondono oltre 700 campioni per riga attiva) non sono emersi vantaggi particolari dall'uso delle suddette strutture e quindi si è preferito optare per una struttura ortogonale, come quella illustrata in figura 1 *a*, che, dal punto di vista dell'organizzazione circuitale, è contraddistinta da una maggior semplicità (bibl. 14, 15).

Nel definire la struttura di campionamento si è altresì valutata l'esigenza di contenere entro limiti ragionevoli il costo e la complessità degli apparati che saranno progettati per funzionare secondo il nuovo standard. In particolare, il trattamento del segnale numerico attraverso le varie fasi della produzione e della post-produzione risulta alquanto semplificato se la struttura di campionamento si ripete da un quadro all'altro; per le stesse ragioni, è inoltre vantaggioso che i due insiemi di campioni rappresentanti le componenti R-Y e B-Y siano co-posizionati tra loro e co-posizionati coi campioni della luminanza (fig. 2).

La possibilità pratica di realizzare i filtri che limitano il contenuto di energia spettrale all'entrata e all'uscita del codec numerico è condizionata dall'esistenza di certo margine di « super-Nyquist » nella frequenza di campionamento; dalle indagini svolte è risultato adeguato un rapporto di circa 2,2 tra frequenza di campionamento e limite superiore (a — 3 dB) della banda base (bibl. 16).

3.3. Qualità globale dell'immagine

Nel definire i parametri dello standard numerico si è perseguito l'obiettivo di garantire un livello qualitativo globale per quanto possibile prossimo a quello dell'immagine originaria, generata nelle sue componenti rosso, verde, blu (RGB). È evidente che, per far fronte alle esigenze della moderna produzione, a cui si è accennato al punto 2, occorre scegliere i parametri di codifica, sia della luminanza che delle componenti di crominanza, tenendo conto dei vari trattamenti a cui il segnale può venir sottoposto lungo la catena.

Nel caso del segnale di luminanza, i problemi maggiori sono connessi con le operazioni che in qualche

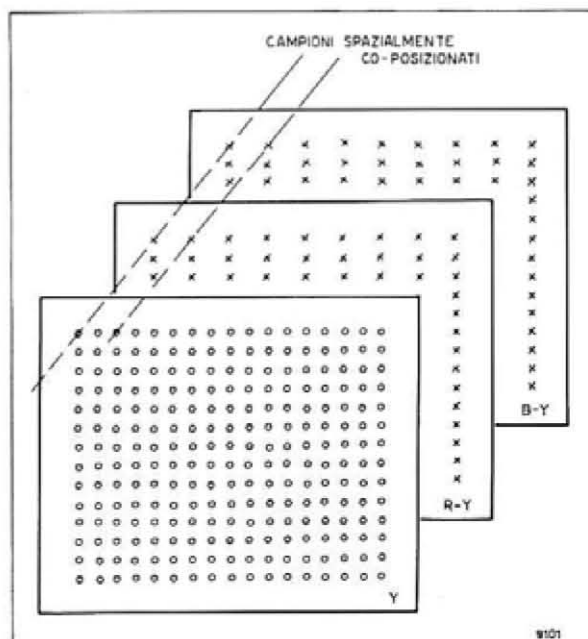



Fig. 2. — Corrispondenza spaziale tra le strutture di campionamento dei segnali Y, R-Y e B-Y nello standard 4 : 2 : 2.

[6] Denominata così in riferimento alla ben nota disposizione dei cinque simboli sulle carte da gioco o sulla faccia dei dadi.

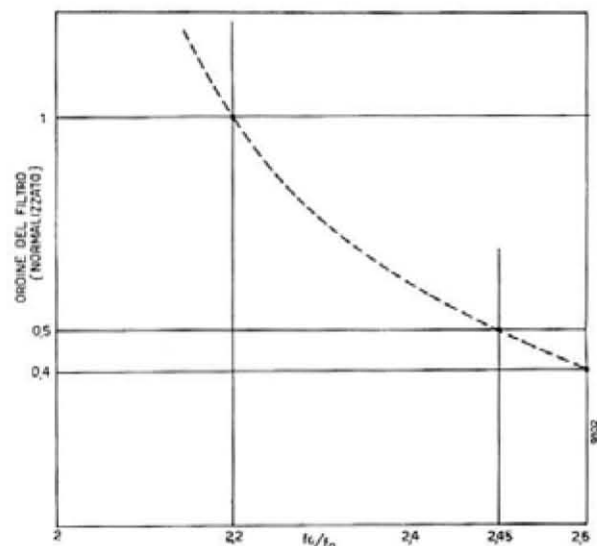



Fig. 3. — Complessità del filtro numerico in relazione al margine di super-Nyquist cioè al rapporto fra la frequenza di campionamento f_c e la banda passante f_o.

modo alterano la geometria dell'immagine; poiché tali operazioni comportano procedimenti di interpolazione nel dominio numerico, è indispensabile l'impiego di filtri numerici la cui realizzazione è governata da esigenze alquanto differenti da quelle che sono alla base del progetto dei filtri analogici impiegati all'entrata e all'uscita del codec numerico.

Per evitare problemi di ritardo di gruppo occorre, fra l'altro, ricorrere a filtri non recursivi, che, come è noto, comportano una quantità maggiore di componenti rispetto a quelli recursivi e quindi un costo più elevato; si è valutato che, a parità di larghezza di banda passante e di attenuazione in banda attenuata, si scenda da 41 celle di ritardo per un filtro FIR (Finite Impulse Response) previsto per funzionare con un margine di super-Nyquist del 9% a 17 celle quando il suddetto margine vale il 27%.

In figura 3 è illustrata, sempre con riferimento ai filtri di tipo FIR, la legge con cui varia l'ordine del filtro in funzione del rapporto f_c/f_o tra la frequenza di campionamento f_c e la banda passante del filtro f_o.

In pratica sono state effettuate prove sperimentali di espansione orizzontale dell'immagine con un filtro variabile avente solo 8 celle di ritardo e con margini di super-Nyquist superiore al 20%; escludendo il comportamento di alcune immagini critiche, il risultato delle prove è stato giudicato soddisfacente [7].

Per quanto concerne le componenti di crominanza, le maggiori problematiche derivano dalle prestazioni richieste al chromakey; tale apparato gioca, infatti, un ruolo determinante nella produzione televisiva e le esigenze da esso imposte in merito al conseguimento di una buona qualità dell'intarsio sono state argomento di approfondita indagine in vari laboratori (bibl. 17, 18, 19).

In figura 4 sono illustrati i risultati di una serie di prove soggettive aventi lo scopo di mettere a confronto varie soluzioni proposte circa lo standard di codifica da studio e, in particolare, la frequenza di campionamento della crominanza (bibl. 20); le terne di valori riportate in ascisse si riferiscono alla frequenza di campionamento, rispettivamente, della luminanza e delle due componenti differenza di colore (per le ragioni esposte al punto 3.2. circa il co-posizionamento dei campioni, le frequenze di campionamento stanno tra loro in rapporti interi).

In ordinate è indicata la variazione di qualità rispetto all'immagine di riferimento; le valutazioni soggettive sono basate sulla scala di qualità CCIR a 5 gradini. Risulta evidente un netto miglioramento qualitativo nel passaggio dalla frequenza di campiona-

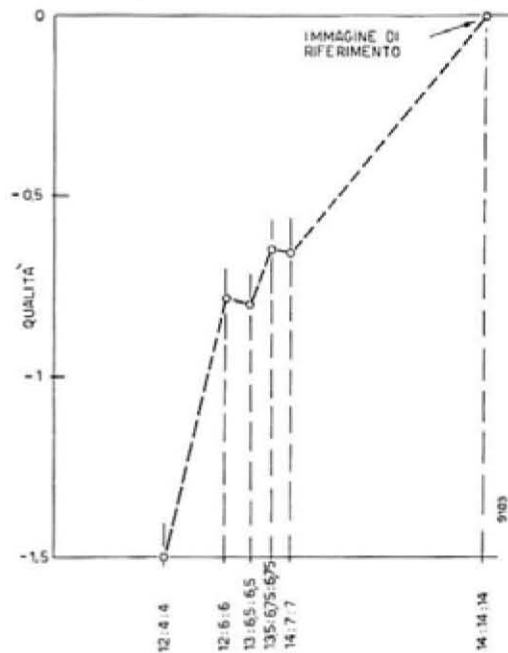



Fig. 4. — Risultati di prove soggettive tendenti a valutare la qualità ottenibile col Chroma-key.

mento della crominanza di 4 MHz a quella di 6 MHz; inoltre, si nota un ulteriore costante incremento di qualità per frequenze di campionamento superiori.

3.4. Numero livelli di quantizzazione e lunghezza della parola

Il rapporto, espresso in dB, tra l'escursione di quantizzazione D (volt) normalmente chiamata «dinamica» e il valore efficace del rumore di quantizzazione N vale:

$$D/N = 10,8 + 6n$$

ove n rappresenta il numero di bit con i quali viene rappresentato ogni campione (bibl. 21).

Tenendo conto che al dinamica del convertitore può essere utilizzata la massimo entro un 90%, per garantire un margine adeguato in caso di sovraoscillazioni dovute ai filtri o comunque di superamento dei livelli video nominali, si può ipotizzare, con un codec funzionante ad 8 bit, un rapporto segnale/rumore di quantizzazione aggirantesi, nei casi peggiori, sui 56 dB; lo standard basato su 8 bit (256 livelli di quantizzazione) è pertanto più che sufficiente ad assicurare che il rumore di quantizzazione resti inferiore alla soglia

[7] Non sono state effettuate prove di compressione per le quali è prevedibile un comportamento leggermente più critico dell'interpolatore.

di visibilità (*), anche nel caso, tutt'altro che improbabile nel periodo di transizione dall'analogico al numerico, in cui siano presenti fino a quattro codecs in cascata lungo la catena (bibl. 22).

Si osservi che lo standard di codifica ad 8 bit è accettabile a livello di interfaccia normalizzata tra apparati. Non è affatto da escludersi che all'interno di un singolo apparato sia necessario operare con più di 8 bit, per evitare processi di arrotondamento suscettibili di produrre un ulteriore incremento del rumore di quantizzazione.

3.5. Registrazione magnetica di segnali numerici

Durante le discussioni che hanno portato alla definizione dello standard, le argomentazioni a favore di un innalzamento della frequenza di campionamento, onde favorire maggiori margini di qualità, vennero seriamente contrastate da considerazioni circa la fattibilità dei registratori videomagnetici numerici operanti ad elevato « bit-rate ».

Per molto tempo sembrò, infatti, che il massimo bit-rate globale registrabile nella pratica, cioè da una macchina affidabile e realizzabile a costi ragionevolmente contenuti, non potesse superare i 140 ÷ 160 Mbit/s. La svolta decisiva si ebbe con lo sviluppo di alcuni prototipi capaci di operare intorno a 230 Mbit/s, e da allora il registratore numerico cessò di rappresentare un fattore cruciale nella scelta dello standard (bibl. 23). Non si possono fare per il momento, previsioni precise circa le specifiche delle macchine che verranno poste sul mercato né circa la data d'immissione, tuttavia, qualche costruttore ipotizza un periodo di sviluppo di circa 3 anni e un formato molto simile agli attuali registratori elicoidali segmentati.

3.6. Manipolazione e immagazzinamento del segnale numerico

I maggiori problemi derivanti dall'aumento della frequenza di campionamento, nella gamma presa in considerazione, consistono nella necessità di ricorrere circuiti integrati funzionanti secondo tecnologie più sofisticate, e pertanto più costosi. Con frequenze di clock superiori ai 12 MHz, occorre infatti sostituire gradualmente i componenti della serie « Low Power Schottky » con componenti della serie FAST e Schottky con un aggravio di costi valutato intorno al 20% nella gamma da 12 MHz a 14 MHz.

Sono state altresì fatte valutazioni circa il costo delle memorie, in particolare le memorie di quadro; un aumento della frequenza di campionamento determina, infatti, in proporzione, un incremento del numero di campioni per unità di tempo da immagazzinare. Grazie all'integrazione su scala via via crescente (*) ed alla organizzazione dei componenti all'interno della memoria di quadro, i costi globali crescono, proporzionalmente, in misura minore rispetto al numero di campioni; si è stimato un incremento di costi del 10% nella gamma da 12 a 14 MHz, tuttavia tale incremento tenderà via via a divenire trascurabile con il progredire dei processi di integrazione VLSI previsti per i prossimi anni.

3.7. Problemi inerenti la distribuzione negli studi e la trasmissione a grandi distanze

In base alle varie proposte avanzate, il bit-rate totale richiesto per trasmettere il segnale numerico nelle tre componenti non risulta superiore in ogni caso ai 225 Mbit/s; tenendo conto di informazioni aggiuntive (bit di parità e sincronizzazione, eventuali segnali audio, segnali ausiliari vari, ecc.) si arriva ad un massimo di 250 Mbit/s.

Studi preliminari hanno indicato che sarà possibile realizzare una distribuzione operante alla suddetta velocità, utilizzando la tecnologia già oggi disponibile; per quanto concerne il costo e l'affidabilità dei multiplatori/demultiplatori da inserire, rispettivamente, all'uscita e all'entrata degli apparati numerici, è stato recentemente creato un gruppo di lavoro dell'UER incaricato di studiare il problema; sulla base di alcune considerazioni preliminari è emerso che la distribuzione totalmente seriale a 250 Mbit dovrebbe risultare più conveniente di altre soluzioni alternative (consistenti ad esempio nel multiplare soltanto i bit di uguale peso delle tre componenti e distribuire il parallelo degli 8 bit multiplati) soprattutto nel caso in cui vengano sviluppati circuiti integrati espressamente per tale uso.

Circa il mezzo di collegamento si prevede di usare il cavo coassiale; sembra, cioè, non necessario far ricorso a fibre ottiche anche se, in previsione di una futura riduzione dei costi dei trasduttori, tale tecnologia offre spunti di un certo interesse.

I problemi della trasmissione sui ponti radio vanno, invece, affrontati tenendo conto delle gerarchie stabilite dal CCITT; nel definire lo standard da studio si è pertanto esplorata, con prove sperimentali, la eventualità di utilizzare il livello a 140 Mbit/s. In particolare si è accertato che, partendo dal più critico (dal punto di vista del « bit-rate ») degli standard proposti, e cioè il 14/7/7, è possibile mettere in atto processi di riduzione del bit-rate, consistenti sostanzialmente nell'espansione della base tempi per occupare gli spazi liberi delle cancellazioni di riga e di trama e nell'impiego di un DPCM (Differential Pulse Code Modulation) ibrido, tali da mantenere la « processabilità » del segnale anche dopo il collegamento; in particolare la qualità del chromakey è risultata praticamente inalterata (bibl. 24).

3.8. Famiglia estensibile di standard di codifica tra loro compatibili

Come è stato ampiamente trattato al punto 2, le moderne tecniche di produzione necessitano che gli impianti siano articolati secondo aree equipaggiate con apparati aventi possibilità operative diverse e, pertanto, suscettibili di fornire prestazioni qualitative diverse.

Nel definire i parametri dello standard principale da studio si è tenuto conto dell'esigenza che esso risulti adeguatamente inserito nel contesto di una famiglia di standard, estensibile sia verso i livelli qualitativi inferiori (per applicazioni tipo ENG) sia verso livelli superiori; quest'ultima soluzione dovrebbe consentire l'effettuazione di produzioni più sofisticate (ad esempio

---

(*) Assunzione valida nell'ipotesi che gli errori di quantizzazione siano distribuiti casualmente. Sono pertanto da escludersi alcune immagini, generate elettronicamente, in cui gli effetti della quantizzazione sono visibili, come pseudocontorni, anche con codifica ad 8 bit.

(*) Tra lo sviluppo delle RAM a 16K e quello della RAM da 64K sono passati meno di sei anni mentre si stanno facendo previsioni circa il passaggio alle 256K entro i prossimi 2 o 3 anni.

nel caso di programmi a definizione più elevata di quella convenzionale per eventuale successivo trasferimento su pellicola cinematografica a 35 mm), oppure dovrebbe rappresentare una porta aperta verso l'alta definizione vera e propria (bibl. 25, 26, 27).

Per consentire di effettuare in modo agevole le operazioni di interpolazione che consentono di convertire da uno all'altro i vari membri della famiglia è indubbiamente conveniente che le frequenze di campionamento stiano tra loro in rapporti semplici e razionali.

Sono stati sperimentati, per il passaggio ai livelli inferiori, rapporti quali 2/1 e 4/3, tuttavia l'indagine non è da ritenersi completa in quanto non è da escludersi che altri valori si prestino meglio ad una ottimizzazione dei risultati ottenibili. Per tale ragione l'attenzione maggiore è stata focalizzata sullo standard principale, che è stato in effetti approvato, rinviando ad un prossimo futuro la discussione sugli altri membri della famiglia.

In base a considerazioni di compromesso tra qualità e contenimento del bit-rate entro termini ragionevoli si è deciso che lo standard da studio sia basato nella forma Y, R-Y, B-Y con R-Y e B-Y aventi metà della frequenza di campionamento della luminanza.

I membri della famiglia sono convenzionalmente indicati tramite una sequenza di numeri interi come 4 : 2 : 2 o 2 : 1 : 1. Tali numeri rappresentano i rapporti tra la frequenza di campionamento della luminanza (primo numero) e le frequenze di campionamento delle componenti differenza di colore (numeri seguenti). Secondo la convenzione adottata il numero 4 rappresenta la frequenza di campionamento della luminanza nello standard da studio.

3.9. Adattabilità dello standard numerico ad entrambe le norme di scansione a 625/50 e 525/60

Tra i criteri seguiti nella definizione dei parametri dello standard, uno dei più importanti è stato indubbiamente l'adattabilità alle due principali norme di scansione tuttora esistenti su base mondiale. Senza entrare nel merito delle approfondite discussioni che hanno portato alla scelta della frequenza di campionamento, ci si limita qui a far osservare che, nella gamma di valori presi in esame tra 12 MHz e 14 MHz, la frequenza di 13,5 MHz è l'unica, che essendo multipla di entrambe le frequenze di riga ([10]), consente un numero intero di campioni per riga nei due sistemi e precisamente 864 e 858.

Nonostante la piccola differenza esistente tra i sopracitati valori, si è voluto ulteriormente aumentare il grado di compatibilità e pertanto è stato introdotto il concetto di «riga attiva numerica» (bibl. 14) la cui durata è leggermente superiore al periodo di riga

Fig. 5. — Relazione temporale tra riga attiva numerica e fronti dei sincronismi di riga analogici.

attiva definito dalle norme sui sistemi analogici; il periodo assegnato alla riga attiva numerica è esattamente lo stesso in entrambi i sistemi a 625 e 525 righe e pertanto contiene lo stesso numero di campioni, con ovvi vantaggi per l'organizzazione delle memorie.

Circa la corrispondenza tra il numero di campioni della riga attiva numerica e la durata dei sincronismi analogici, in figura 5 è illustrata la soluzione proposta dall'EBU per i sistemi di scansione a 625 righe.

4. Specifiche del sistema di codifica per gli studi.

Vengono qui illustrati i vari punti della Raccomandazione AA/11 del CCIR, in cui sono indicate le specifiche di base del sistema di codifica numerica per gli studi; le specifiche sono valide per impianti operanti sia secondo le norme di scansione a 625 righe sia secondo quelle a 525 righe.

Il testo della suddetta raccomandazione è articolato in quattro punti, i cui primi tre riguardano l'intero sistema ed in particolare trattano la forma del segnale codificato, il concetto di famiglia estensibile di standard compatibili e la struttura di campionamento, mentre il quarto si riferisce dettagliatamente allo standard 4 : 2 : 2, cioè al livello principale della famiglia di standard compatibili.

4.1. Forma del segnale codificato

Il segnale codificato deve essere nel formato in componenti (segnale luminanza e segnali differenza di colore oppure segnali rosso, verde, blu).

4.2. Famiglia estensibile di standard di codifica mutuamente compatibili

Le specifiche di codifica debbono consentire le costituzione di una famiglia di standard tra loro compatibili; la famiglia deve essere estensibile ed evolutiva nel senso che nuovi membri possono essere definiti in futuro.

Deve risultare possibile interfacciare in modo semplice due qualsiasi dei membri della famiglia.

Il membro della famiglia da impiegare come interfaccia numerica tra gli apparati funzionanti al livello qualitativo «da studio» e per lo scambio internazionale dei programmi ([11]) deve essere quello in cui le frequenze di campionamento della luminanza e dei segnali differenza di colore stanno fra loro in rapporto 4 : 2 : 2.

In un eventuale membro della famiglia funzionante al livello gerarchico immediatamente superiore a quello

---

([10]) Rispettivamente 15625 Hz per i sistemi a 625 righe e 15734,265 Hz per i sistemi a 525 righe. In particolare, si ha 864 · 15625 = 858 · 15734,265.

([11]) È uno dei punti della raccomandazione che hanno suscitato il maggior numero di controversie. L'obiettivo principale della discussione era infatti quello di fissare un riferimento operativo ad almeno uno dei membri della famiglia, ed in particolare a quello su cui era stata concentrata la maggior quantità di indagini sperimentali. In pratica, lo standard 4 : 2 : 2 è quello riservato alla normale produzione televisiva e costituisce l'interfaccia tra lo studio e il sistema di trasmissione.

Fig. 6. — Corrispondenze tra livelli dei segnali analogici e livelli di quantizzazione.

a) Segnale di luminanza; la dinamica di utilizzazione del convertitore si estende dal livello 16 compreso, al livello 235 compreso, per un totale di 220 livelli; rimangono rispettivamente 16 livelli (dallo 0 compreso, al 15 compreso) e 20 livelli (dal 236 compreso, al 255 compreso) per assicurare un margine adeguato alle sovraoscillazioni del segnale analogico. b) Segnale differenza di colore; la dinamica si estende dal livello 16 compreso, al livello 239 compreso, per un totale di 224 livelli; rimangono 16 livelli di margine ai due estremi.

previsto per la normale produzione, si prevede la possibilità di operare sia in Y, R-Y, B-Y che in R, G, B; le frequenze di campionamento della luminanza e dei segnali differenza di colore (oppure delle componenti R, G, B) dovrebbero stare tra loro in rapporto 4 : 4 : 4.

Nella tabella 2 sono illustrati i parametri, indicati al momento solo a titolo di proposta, dello standard 4 : 4 : 4.

4.3. SPECIFICHE RIGUARDANTI LE STRUTTURE DI CAMPIONAMENTO (VALIDE PER TUTTI I MEMBRI DELLA FAMIGLIA)

4.3.1. Le strutture di campionamento debbono risultare statiche sul piano dell'immagine.

4.3.2. Nei casi in cui i campioni rappresentano il segnale nella forma di componente di luminanza e di differenza di colore prese simultaneamente ([12]), ogni coppia di campioni relativi ai segnali differenza di colore debbono essere co-posizionati rispetto ai corrispondenti elementi dell'immagine. Se i campioni rappresentano il segnale della forma di componenti R, G, B, essi debbono, altresì essere co-posizionati.

([12]) Non riguarda, ad esempio, il caso prevedibile nei livelli inferiori, in cui i segnali R-Y e B-Y vengano eventualmente trasmessi sequenzialmente (tipo SECAM).

4.3.3. Al fine di assicurare, su scala mondiale, l'adattabilità di ogni membro della famiglia agli impianti funzionanti con le attuali norme di scansione, il numero di campioni per riga, stabiliti, rispettivamente per i sistemi a 625 e 525 righe, deve risultare compatibile (preferibilmente si dovrà fissare lo stesso numero di campioni per riga).

4.4. SPECIFICHE DELLO STANDARD 4 : 2 : 2

In tabella 1 sono riportate, in dettaglio, le specifiche dello standard 4 : 2 : 2 da impiegare nelle interfacce

Tabella 1

SPECIFICHE DELLO STANDARD 4:2:2

Parametri	Sistemi a 525 righe, 60 trame/s	Sistemi a 625 righe, 50 trame/s
1. Segnali codificati	Y, R-Y, B-Y (*)	
2. Numero di campioni per riga: — segnale di lumin. (Y) — segnali differenza di colore (R-Y, B-Y)	 858 429	 864 432
3. Struttura di campionamento	Ortogonale, ripetitiva di riga, trama e quadro. Campioni di R-Y e B-Y composizionati ad ogni riga con i campioni dispari (1ª, 3ª, 5ª, ecc.) della Y	
4. Frequenza di campionamento: segnale di luminanza segnali differenza di colore	 13,5 MHz 6,75 MHz	
5. Tipo di codifica	PCM a distribuzione uniforme dei livelli di quantizzazione; 8 bit per campione, sia per la luminanza che per i segnali differenza di colore	
6. Numero di campioni per riga attiva numerica: — segnale di luminanza — segnale differenza di colore	 720 360	
7. Corrispondenza tra livelli del segnale analogico e livelli di quantizzazione (**) — segnale di luminanza — segnali differenza di colore	220 livelli di quantizzazione col livello del nero corrispondente al livello 16 e il livello di picco del bianco corrispondente al livello 235; 224 livelli di quantizzazione con lo zero corrispondente al livello 128	

(*) Per la definizione dei segnali R-Y e B-Y vedere Appendice.

(**) Vedere anche figura 6. La dinamica utile del convertitore viene limitata a 220 livelli per il segnale di luminanza e a 224 livelli per i segnali differenza di colore, lasciando pertanto adeguati margini per le sovraoscillazioni dei segnali analogici o per eventuali errori nella regolazione del guadagno sulla tratta analogica. Nel caso della luminanza il margine superiore è stato tenuto più ampio (20 livelli) poichè si è valutato che le conseguenze negative delle limitazioni di ampiezza sono più fastidiose in corrispondenza dei bianchi.

Tabella 2

Specifiche proposte per l'eventuale standard 4:4:4

Parametri	Sistema a 525 righe 60 trame/s	Sistema a 625 righe, 50 trame/s
1. Segnali codificati	R, G, B, oppure Y, R-Y (*), B-Y (*)	
2. Numero di campioni per riga (riferito a ciascuno dei 3 segnali)	858	864
3. Struttura di campionamento	Ortogonale, ripetitiva di riga, trama e quadro. Le 3 strutture di campionamento debbono coincidere tra loro e inoltre debbono coincidere con la struttura di campionamento della luminanza nello standard 4:2:2	
4. Frequenza di campionamento dei 3 segnali	13,5 MHz	
5. Tipo di codifica	PCM a distribuzione uniforme dei livelli di quantizzazione. Almeno 8 bit per campione	
6. Durata della riga attiva numerica espressa in numero di campioni	Almeno 720	
7. Corrispondenza tra livelli del segnale analogico e livelli di quantizzazione		
RGB oppure Y (**)	220 livelli di quantizzazione col livello del nero corrispondente al livello 16 e il livello di picco del bianco corrispondente al livello 235	
Segnali differenza di colore (**)	224 livelli di quantizzazione con lo zero corrispondente al livello 128	

(*) Si intendono segnali differenza di colore a piena banda come il segnale di luminanza Y.
(**) Se viene usata la versione Y, R-Y, B-Y.

tra apparati operanti a livello qualitativo da studio e nello scambio internazionale dei programmi.

5. Altri parametri relativi allo standard 4 : 2 : 2 proposti per ulteriori indagini.

Nel testo della Raccomandazione AA/11 compaiono, ovviamente, soltanto le specifiche del sistema sulle quali è stata raggiunta l'unanimità dell'accordo. Tuttavia, tra le varie proposte presentate al CCIR, figurano altri parametri (bibl. 8), che non hanno conseguito l'approvazione poiché si è ritenuto utile riservare ad essi un ulteriore periodo di riflessione, ma che hanno discrete probabilità di essere inclusi nel testo della Raccomandazione durante il prossimo periodo di studi (1982-1986) qualora indagini più approfondite ne confermino la validità. A titolo di informazione, le suddette proposte sono riportate, in sintesi, in tabella 3.

Tabella 3

Larghezza di banda dei segnali codificati	— Canale di luminanza: nominalmente piatto fino ad almeno 5.5 MHz e attenuato di almeno 12dB a 6.75 MHz — Canali differenza di colore: nominalmente piatti fino ad almeno 2.75 MHz e attenuati di almeno 12 dB a 3.375 MHz
Caratteristiche dei filtri analogici impiegati nei codecs	Attenuazione in banda oscura dei filtri impiegati nella conversione analogico-numerica: — maggiore di 45 dB a 8 MHz per il segnale di luminanza — maggiore di 45 dB a 4 MHz per i segnali differenza di colore Il contenuto spettrale dei segnali differenza di colore dovrà essere sagomato da filtri a taglio morbido inseriti in eventuali codificatori a colori di tipo analogico. L'equalizzazione di tipo $x/\sin x$, qualora richiesta, deve essere inserita soltanto alla fine della tratta numerica, al momento della conversione numerico-analogica.
Relazione temporale tra riga attiva numerica e forme d'onda analogiche dei segnali di sincronismo	v. fig. 5

6. Conclusioni.

Dopo un lungo periodo contraddistinto dall'esistenza di una vasta gamma di standard differenti all'interno degli impianti di produzione, la televisione si accinge ad entrare nell'era delle tecniche numeriche cancellando, in un solo passo, la maggior parte delle differenziazioni ereditate dal passato.

La codifica del segnale a colori nella forma delle tre componenti separate non reca più alcuna traccia degli attuali sistemi PAL, SECAM o NTSC; inoltre essa consente di produrre programmi adottando tecniche assai più sofisticate e ottenendo un livello qualitativo superiore che con gli impianti analogici.

Il conseguimento, su base mondiale, di un accordo sui principali parametri dello standard di codifica, e, in particolare, circa il valore delle frequenze di campionamento delle tre componenti (rispettivamente 13,5, 6,75 e 6,75 MHz), consente, di fatto, di adattare gli stessi impianti numerici sia ai paesi operanti secondo norme di scansione a 625 righe sia in quelli operanti a 525 righe.

I benefici di tale accordo sono stati valutati in termini sia economici che operativi. Secondo una previsione effettuata da esperti dell'UER (bibl. 14) il risparmio sul costo degli apparati derivante dalla loro progettazione in versione unica si aggirerebbe intorno al 10 ÷ 15%; inoltre, sono state svolte considerazioni riguardanti la semplificazione nello scambio internazionale dei programmi e le potenzialità di uno sviluppo coerente, su scala mondiale, dei sistemi di comunicazione basati sulla diffusione di immagini.

Come è stato ripetutamente sottolineato nel presente lavoro, il sistema di codifica numerica che costi-

tuisce oggetto della Raccomandazione AA/11 riguarda esclusivamente gli apparati operanti all'interno degli studi televisivi; resta, pertanto, inteso che all'atto della messa in onda del programma in una determinata area di servizio, il segnale necessita di essere convertito nello standard di trasmissione ivi operante (PAL, SECAM o NTSC).

Sarà comunque materia di indagine per i prossimi anni, l'individuazione di possibili soluzioni tecniche che consentano di portare fino a casa dell'utente lo stesso livello qualitativo oggi conseguibile nello studio numerico; nel perseguire tale obiettivo si adotterà il nuovo standard come riferimento per la conversione tecnologica dell'intera catena televisiva, dalla telecamera al ricevitore domestico, anziché utilizzarlo unicamente come strumento di produzione dei programmi.

APPENDICE

DEFINIZIONE DEI SEGNALI DI LUMINANZA (Y) E DIFFERENZA DI COLORE (R-Y, B-Y)

Con riferimento al Rapporto 624-1 del CCIR, ed indicando, per semplicità, con R, G, B i segnali primari precorretti di γ, si ha:

$$\begin{aligned} Y &= 0{,}299\ R + 0{,}587\ G + 0{,}114\ B \\ R\text{-}Y &= 0{,}701\ R - 0{,}587\ G - 0{,}114\ B \\ B\text{-}Y &= -0{,}299\ R - 0{,}587\ G + 0{,}886\ B. \end{aligned}$$

Supponendo che l'ampiezza dei primari sia normalizzata all'unità (cioè regolando il livello massimo al valore di 1,0 volt), i valori ottenuti rispettivamente per il bianco, il nero e i colori primari saturi nonché i complementari sono quelli riportati in tabella A.

TABELLA A

	R	G	B	Y	R-Y	B-Y
Bianco	1.0	1.0	1.0	1.0	0	0
Nero	0	0	0	0	0	0
Rosso	1.0	0	0	0.299	0.701	— 0.299
Verde	0	1.0	0	0.587	— 0.587	— 0.587
Blu	0	0	1.0	0.114	— 0.114	0.886
Giallo	1.0	1.0	0	0.886	0.114	— 0.886
Ciano	0	1.0	1.0	0.701	— 0.701	0.299
Magenta	1.0	0	1.0	0.413	0.587	0.587

Mentre la dinamica del segnale Y si estende da 1,0 a 0, quella dei segnali R-Y e B-Y si estende invece, rispettivamente da 0,701 a — 0,701 e da 0,886 a — 0,886. Per riportare all'unità l'escursione dei segnali differenza di colore (es. da 0,5 a — 0,5) occorre alterarne il valore moltiplicandolo per i seguenti coefficienti:

$$\alpha_R = \frac{0{,}5}{0{,}701} = 0{,}713; \qquad \alpha_B = \frac{0{,}5}{0{,}886} = 0{,}564$$

Si ha pertanto:

$$\begin{aligned} (R\text{-}Y)^* &= 0{,}713(R\text{-}Y) = 0{,}500\ R - 0{,}419\ G - 0{,}081\ B \\ (B\text{-}Y)^* &= 0{,}564(B\text{-}Y) = -0{,}169\ R - 0{,}331\ G + 0{,}500\ B \end{aligned}$$

ove $(R\text{-}Y)^*$ e $(B\text{-}Y)^*$ sono i valori ri-normalizzati dei segnali differenza di colore.

Nel caso della codifica numerica a 8 bit con scala di quantizzazione lineare si hanno disponibili 256 livelli; la gamma di numeri binari generati dal processo di codifica si estende da 00000000 a 11111111 mentre, per i corrispondenti valori decimali, si estende da 0 a 255.

Assegnando 220 livelli (dei 256 disponibili) all'escursione del segnale di luminanza (per garantire un margine operativo) e facendo corrispondere il nero al livello 16, il valore decimale $\overline{Y}$ del livello corrispondente ad un generico valore Y della luminanza è:

$$\overline{Y} = <220\ Y + 16>$$

ove il simbolo < > indica, come noto, il numero intero più prossimo.

Analogamente, assegnando ai segnali differenza di colore 224 livelli e facendo corrispondere il valore nullo al valore 128, i valori decimali $\overline{(R\text{-}Y)}$ e $\overline{(B\text{-}Y)}$ dei livelli corrispondenti a generici valori (R-Y) e (B-Y) delle componenti normalizzate sono:

$$\begin{aligned} \overline{(R\text{-}Y)} &= <224 \cdot 0{,}713\ (R\text{-}Y) + 128> \\ \overline{(B\text{-}Y)} &= <224 \cdot 0{,}564\ (B\text{-}Y) + 128> \end{aligned}$$

cioè, in definitiva:

$$\begin{aligned} \overline{(R\text{-}Y)} &= <160\ (R\text{-}Y) + 128>, \\ \overline{(B\text{-}Y)} &= <126\ (B\text{-}Y) + 128>. \end{aligned}$$

(2899)

BIBLIOGRAFIA

1. - Draft Recommandation AA/11. CCIR - XV Plenary Assembly - Geneva, 1982 - Doc. 11/1027 - rev. 1.
2. - GUINET Y.: *Evolution of the EBU's position in respect of the digital coding of television.* « EBU Review », n. 187, giugno 1981.
3. - CCIR (1978-1982) - Doc. 11/323 - Italy.
4. - AKRICH C., ZACCARIAN, P.: *Production requirements for digital television systems.* « EBU Review », n. 187, giugno 1981.
5. - CONNOLLY W. G.: *Electronic production and post-production of television programmes using single camera technique.* « 12° Symposium Internazionale di Televisione ». Montreux, giugno 1981.
6. - LOWRY J. D.: *Digital processing - An overview.* « SMPTE - 14th Annual Conference in Toronto », febbraio 1980.
7. - JONES A. H.: *Digital video coding: the selection of parameters to meet requirements foreseen.* « EBU Review », n. 187, giugno 1981.
8. - Draft Rep. 629-1 (MOD F) - CCIR (1978-1982) - Doc. 11/5048.

9. - Draft Rep. AG/11 - CCIR (1978-1982) - Doc. 11/5041.

10. - CCIR (1978-1982) - Doc. 11/14 - (EBU).

11. - CCIR (1978-1982) - Doc. 11/15 - (EBU).

12. - SABATIER J., KRETZ F.: *Sampling the components of 625 line colour television signals.* « EBU Review », n. 171, ottobre 1978.

13. - TONGE G. J.: *The sampling of television images - IBA Experimental and Development.* « Report 112/81 ».

14. - CCIR (1978-1982) - Doc. 11/330 - (EBU).

15. - CCIR (1978-1982) - Doc. 11/89 - (Poland).

16. - CCIR (1978-1982) - Doc. 11/327 - (Italy).

17. - CCIR (1978-1982) - Doc. 11/261 - (France).

18. - CCIR (1978-1982) - Doc. 11/285 - (United Kingdom).

19. - LUCAS K., NASSE D.: *Case studies in the implementation on chroma-key.* « EBU Review », n. 187, giugno 1981.

20. - DARBY P. J.: *Subjective assessment of television picture quality.* « IBA Experimental and Development », Report 114/81.

21. - BARBIERI G. F., COMINETTI M., D'AMATO P.: *La codifica numerica lineare del segnale video.* « Elettronica e Telecomunicazioni », n. 3, 1972.

22. - ALBERICO F., BARBIERI G. F., D'AMATO P., GHIDINI I.: *Determinazione soggettiva dei parametri di un sistema PCM - video lineare.* « Elettronica e telecomunicazioni », n. 3, 1972.

23. - DRURY G. M.: *A digital video tape recorder for a component standard requiring 216 ÷ 228 Mbit/s.* « IBA Experimental and Development », Interim, Report 110/81.

24. - WILSON E. J.: *Bit-rate reduction to 140 Mbit/s from a television component standard requiring 216÷228 Mbit/s.* « IBA Experimental and Development », Interim Report 111/81.

25. - CCIR (1978-1982) - Doc. 11/31 - (USA).

26. - CCIR (1978-1982) - Doc. 11/16 - (EBU).

27. - CCIR (1978-1982) - Doc. 11/114 - (France).

Le origini del video digitale (La raccomandazione ITU-R BT.601)

Testo e figure tratte dall'articolo di Marzio Barbero e Natasha Shpuza, Elettronica e Telecomunicazioni, Aprile 2003. Il testo è aggiornato sulla base della attuale versione (BT.601-6, 2007) della raccomandazione.

1. Premessa

Alla base della rapida evoluzione dei sistemi di codifica digitale dell'informazione video sono le specifiche relative al campionamento e alla codifica individuate dalla Raccomandazione ITU-R BT.601. L'articolo [1], riprodotto come primo capitolo, forniva una completa e competente analisi della nuova raccomandazione: la prima versione risale infatti a tale anno. Questo capitolo intende riproporre brevemente le caratteristiche principali di queste specifiche poiché esse spesso hanno implicazioni rilevanti nella definizione dei sistemi di compressione del segnale video e sulle tecniche di editing. La definizione della Rac. ITU-R BT.601 è l'evento alla base dello sviluppo e alla diffusione degli apparati di ripresa, di editing e di distribuzione.

2. Cenni storici

Negli anni '70, le tecnologie numeriche si erano sufficientemente sviluppate da consentire all'industria di produrre i primi apparati digitali in grado di operare su segnali video di qualità adatta alla realizzazione di prodotti televisivi: nel 1981 l'UER e SMPTE riuscirono a definire i parametri essenziali per raggiungere un accordo a livello mondiale. In particolare i ricercatori della Rai contribuirono, così come quelli dei maggiori laboratori e centri di ricerca degli enti televisivi, alla sperimentazione tecnica, nell'ambito dei gruppi di lavoro UER e del ITU, necessaria alla definizione di tali parametri.

La prima versione delle specifiche si limitava alla definizione dei parametri relativi al primo membro di una "famiglia estensibile di standard per la codifica digitale compatibile" basata sull'adozione di una frequenza di campionamento unica, pari a 13,5 MHz, per i formati di immagine 4:3 a 625 righe/50 Hz e 525 righe/60 Hz. Successivamente furono previste membri della famiglia con un formato di immagine 16:9 e una frequenza di campionamento pari a 18 MHz, con un conseguente aumento numero di campioni per riga, al fine di mantenere costante la risoluzione orizzontale. Tali membri non trovarono di fatto applicazione pratica e nella versione attualmente in vigore, la 601-6 del 2007 [2], l'unica frequenza prevista torna ad essere quella pari a 13,5 MHz, sia per immagini 4:3, sia per quelle 16:9, con la motivazione che le prestazioni offerte sono adeguate per gli attuali sistemi di trasmissione.

3. Motivazioni alla base dello standard

La decisione di passare da una codifica del segnale video di tipo analogico (NTSC, PAL, SECAM) ad una codifica di tipo numerico ha varie motivazioni:

- Consentire l'uso di reti digitali per trasportare il segnale: i collegamenti numerici si possono considerare "trasparenti" al tipo di informazione trasportata (ad esempio: dati, audio, video) e al formato di codifica dell'informazione (ad esempio, nel caso del video: formati a componenti o compositi, a 625 o 525 righe)

- Utilizzare apparati in grado di elaborare l'informazione video in modo da consentire l'ideazione di programmi e l'utilizzazione di linguaggi tecnico-espressivi (ad esempio effetti speciali e studio virtuale) assolutamente non pensabili nel caso si fosse continuato ad operare nel dominio analogico.

La raccomandazione è stata definita con lo scopo di avere la maggior parte dei parametri in comune nel caso dei formati a 625 e 525 righe al fine di consentire economie di scala nella realizzazione degli apparati e di facilitare lo scambio internazionale dei programmi. In particolare si è scelto di avere frequenze di campionamento uguali (per formati a 525 e 625 righe) e una codifica basata su tre componenti.

4. La famiglia estensibile di standard compatibili per la codifica digitale

4.1 4:4:4 e 4:2:2

I segnali possono essere sotto forma RGB, cioè corrispondenti ai tre colori primari (rosso, verde e blu, ovvero RGB *red-green-blue*). La colorimetria (le coordinate colorimetriche CIE) sono quelle specificate per i sistemi analogici a 625 e 525 righe, ma, con l'attuale versione [2], è consentito l'uso della colorimetria e dei parametri per la matrice colorimetrica definita dall Rac. ITU-R BT.1361[3], al fine di permettere la compatibilità diretta con i sistemi HDTV [4] (figura 8) .

In questo caso tutte le tre componenti sono campionate alla stessa frequenza: i membri della famiglia che rispondono a queste caratteristiche sono denominati 4:4:4. La struttura di campionamento prevede che i campioni relativi alle tre componenti siano co-posizionati (figura 1).

A partire dai segnali elettrici precorretti di gamma E'_R, E'_G e E'_B, all'uscita dei sensori della telecamera è possibile ottenere come loro combinazione lineare altri tre segnali digitali denominati Y (luminanza), C_R e C_B (differenza-colore), anche in questo caso la struttura è 4:4:4, cioè le tre componenti sono caratterizzate dalla stessa frequenza di campionamento e sono co-posizionati.

A partire dalle componenti Y, C_R e C_B nel formato 4:4:4 è possibile ottenere il formato 4:2:2, caratterizzato dal fatto che le due componenti differenza-colore sono limitate in banda e campionate a metà della frequenza con cui è campionata la luminanza.

Ovviamente la limitazione di banda implica che i formati 4:2:2 siano meno adatti ad applicazioni in cui si richieda una elevata qualità anche in presenza di numerose e complesse elaborazioni del segnale video.

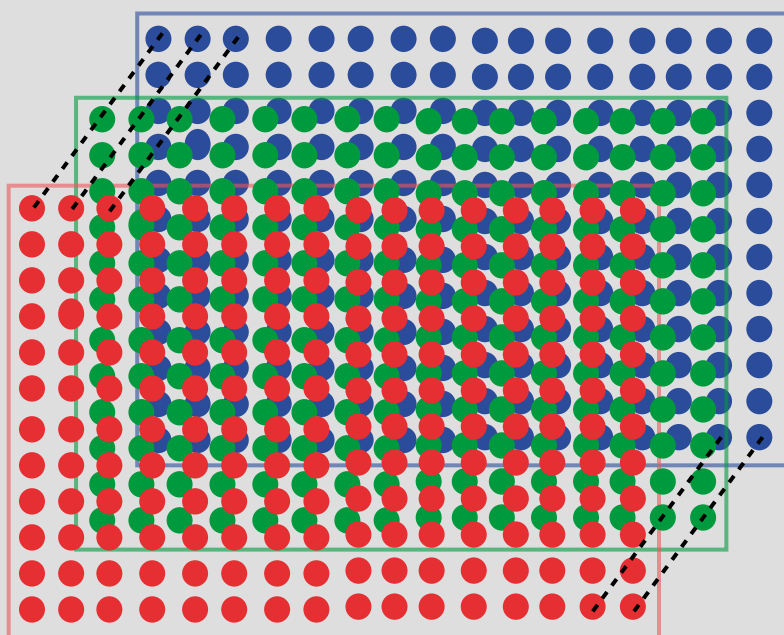



Fig. 1 - Nel caso di struttura di campionamento 4:4:4, le tre componenti sono campionate alla stessa frequenza e i tre campioni corrispondenti allo stesso elemento di immagine sono co-posizionati: ad esempio, nel caso di RGB i campioni rosso, verde e blu hanno lo stesso reticolo di campionamento.

D'altro canto la scelta di prevedere il formato 4:2:2 è essenzialmente economica: si riduce del 25% il numero di campioni da memorizzare o trasmettere. Inoltre sia i sistemi per la diffusione, sia quelli analogici (PAL, SECAM, NTSC) che quelli digitali compressi secondo lo standard MPEG (DVB, DVD) sono caratterizzati da forte limitazione di banda delle componenti di colore e quindi non vi sono degradamenti nel passaggio dal formato professionale a quello di diffusione.

Quando vennero definiti i parametri di codifica, molta attenzione fu fatta sulla scelta della struttura di campionamento e sulla banda necessaria per le componenti, per assicurare buone prestazioni anche nel caso di operazioni di postproduzione complesse, quali il chroma-key o intarsio. Particolare cura è posta nelle specifiche per i filtri a cui sottoporre i segnali Y o R,G e B (figura 2) e i segnali differenza- colore sia nel caso siano ottenuti direttamente a partire dai segnali analogici (figura 3), sia

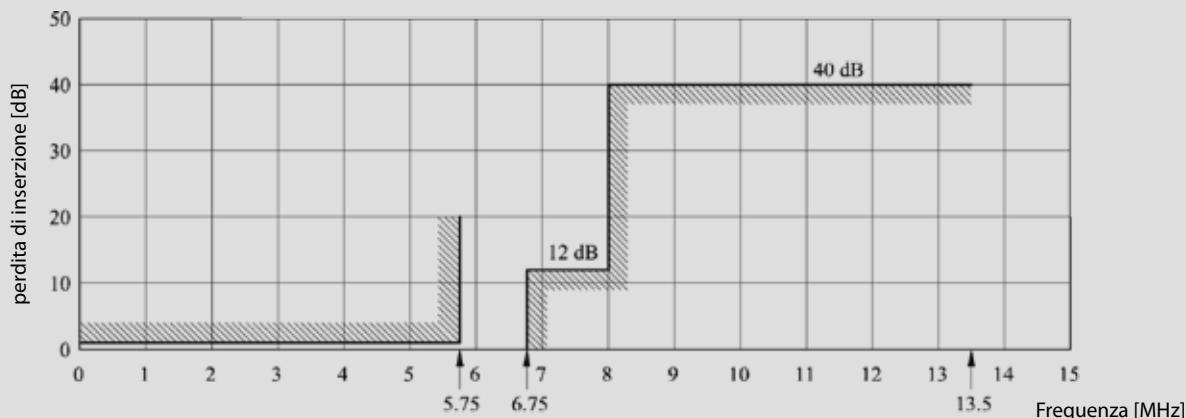



Fig. 2 - Specifiche per un filtro per i segnali di luminanza o RGB quando si utilizza la frequenza di campionamento 13,5 MHz: maschera per la caratteristica perdita d'inserzione in funzione della frequenza.

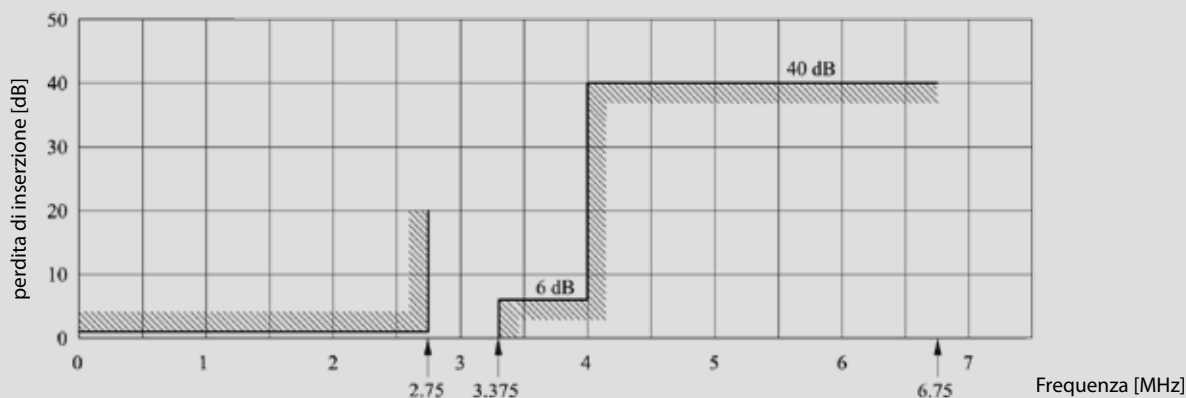



Fig. 3 - Specifiche per un filtro per i segnali di differenza-colore quando si utilizza la frequenza di campionamento 6,75 MHz: maschera per la caratteristica perdita d'inserzione in funzione della frequenza.

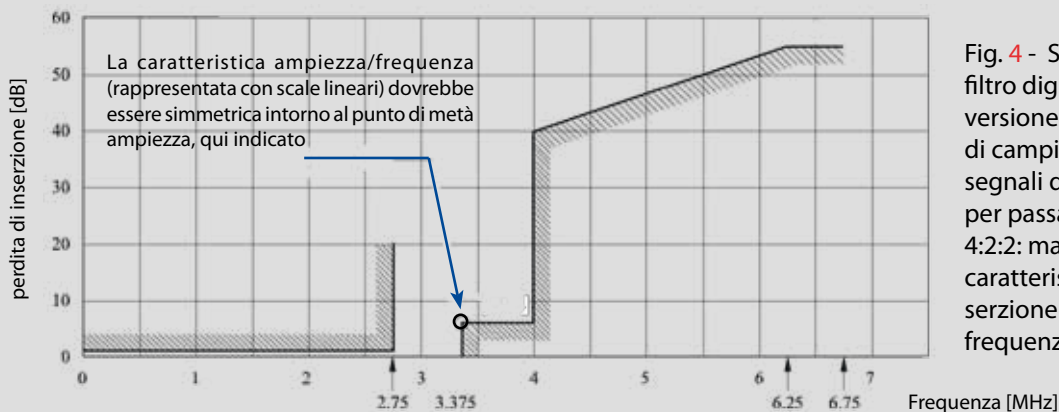



Fig. 4 - Specifiche per un filtro digitale per la conversione dalla frequenza di campionamento dei segnali differenza-colore per passare da 4:4:4 a 4:2:2: maschera per la caratteristica perdita d'inserzione in funzione della frequenza.

a partire dai segnali digitali mediante filtri numerici (figura 4).

Nel caso di struttura di campionamento 4:2:2 ciascuna coppia relativa ai due segnali differenze-colore deve essere co-posizionata spazialmente con i campioni di posizione dispari della luminanza (figura 5).

4.2 La quantizzazione

Le parole digitali che rappresentano i valori dei campioni possono essere a 8 bit o a 10 bit, gli otto bit più significativi sono la parte intera, mentre i restanti due bit, se presenti, sono da considerare parte frazionaria e se non sono indicati sono da supporre uguali alle cifre binarie 00.

Con 8 bit sono disponibili 256 livelli di quantizzazione equispaziati (da 0000 0000 a 1111 1111 in notazione binaria, o da 0 a 255 in notazione decimale). Le configurazioni 0 e 255 sono riservate per la sincronizzazione, mentre quelle da 1 a 254 sono destinate al video.

Per consentire opportuni margini operativi (ad esempio per evitare che calcoli per realizzare filtri digitali generino valori video non rappresentabili correttamente) l'informazione di luminanza occupa solo 220 dei livelli disponibili: il nero corrisponde al livello 16 e il bianco al livello 235. Analogamente i segnali differenze di colore devono occupare solo 225 livelli e il valore 0 di ciascun segnale differenza di colore deve corrispondere al livello digitale 128.

4.3 I membri della famiglia

La tabella 1 riporta i parametri per il membro 4:2:2 e frequenza di campionamento 13,5 MHz della famiglia di standard. Questi parametri si applicano sia nel caso in cui il formato d'immagine sia 4:3 sia nel caso in cui sia 16:9 (figura 6).

La scelta della stessa frequenza di campionamento per i sistemi a 625 righe e quelli a 525 righe facilita la realizzazione dei sistemi di instradamento e commutazione del segnale televisivo numerico negli studi e centri di produzione televisivi: il flusso binario complessivo è pari, in entrambi i casi, a 270 Mbps (l'interfaccia di interconnessione è specificata da una successiva raccomandazione [5], la cui ultima versione è del 2007, dove i campioni di luminanza e di differenza-colore sono multiplati e a ciascun campione corrispondono 10 bit).

Si osservi che il numero di campioni costituenti la riga attiva è uguale per i sistemi a 625 righe e quelli a 525 righe (720 campioni Y), malgrado differiscano le durate totali delle righe (sono rispettivamente 864 e 858). Questa scelta ha consentito di facilitare la struttura dei sistemi in grado di elaborare o memorizzare il segnale video digitale. Non solo la quantità di dati per riga attiva è costante, ma anche la quantità di dati utili (parte attiva del segnale video) complessiva è costante: infatti, anche se questo parametro non è indicato nella [2], il numero di righe attive normalmente elaborato o memorizzato è

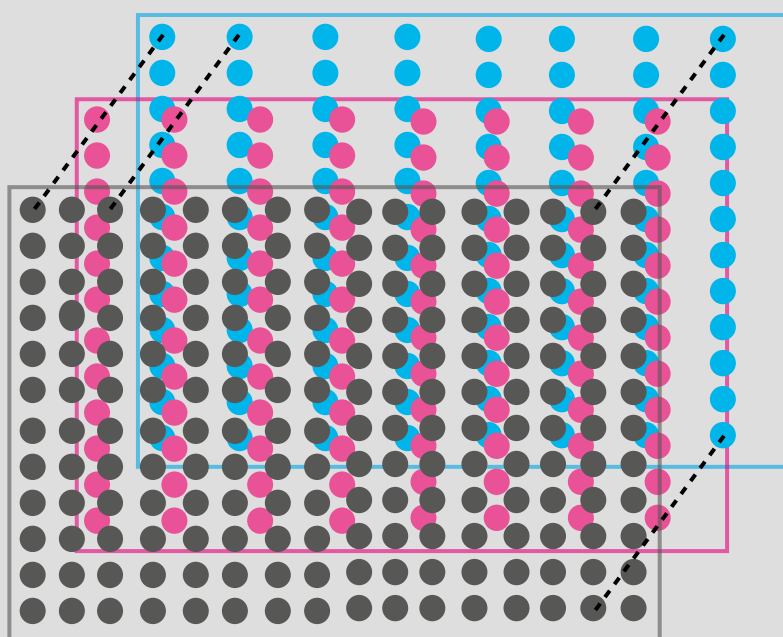



Fig. 5 - Nel caso di struttura di campionamento 4:2:2, il numero di campioni relativi alle componenti differenza-colore è metà rispetto ai campioni di luminanza. Entrambi i campioni differenza-colore sono co-posizionati con i campioni di ordine dispari relativi alla luminanza.

rispettivamente 576 nel caso di sistemi a 25 quadri al secondo e 480 nel caso di quelli a 30 quadri al secondo e pertanto il numero totale di elementi d'immagine risulta essere in entrambi i casi 10 368 000 al secondo.

Nel caso del membro 4:4:4 (utilizzabile sia per R,G, B che per Y, C_R e C_B) valgono parametri del tutto simili a quelli indicati nella tabella precedentemente citata, tranne che per tutte le tre componenti valgono la stessa frequenza di campionamento (13,5 MHz) e lo stesso numero di campioni per riga totale e per riga attiva (720 per ciascuna componente).

Parametri	Sistemi a 525 righe e 60 semiquadri	Sistemi a 625 righe e 50 semiquadri
Segnali codificati Y, C_R , C_B	Questi segnali sono ottenuti a partire dai segnali precorretti di gamma, vale a dire E'_Y, E'_R-E'_Y, E'_B-E'_Y	
Numero di campioni per l'intera riga - segnale di luminanza (Y) - ciascun segnale differenza-colore (C_R, C_B)	 858 429	 864 432
Struttura di campionamento	Ortogonale, ripetitiva a livello di riga, semiquadro e quadro. I campioni C_R, C_B sono coposizionati rispetto ai campioni dispari (1°, 3°, 5° …) in ciascuna riga	
Frequenza di campionamento - segnale di luminanza - ciascun segnale differenza di colore	 13,5 MHz 6,75 MHz La tolleranza per la frequenza di campionamento deve coincidere con la tolleranza per la frequenza di riga dello standard televisivo a colori corrispondente	
Formato di codifica	PCM uniformemente quantizzato, 8 o 10 bit per campione, per il segnale di luminanza e per ciascuno dei segnali differenza di colore	
Numero di campioni per riga attiva digitale - segnale di luminanza - ciascun segnale differenza di colore	 720 360	
Relazione temporale orizzontale analogico-digitale - dalla fine della riga attiva digitale a O_H	16 periodi di clock per la luminanza	12 periodi di clock per la luminanza
Corrispondenza tra i livelli del segnale video e i livelli di quantizzazione - estensione della scala di quantizzazione - segnale di luminanza - ciascun segnale differenza di colore	 Da 0 a 255 220 livelli di quantizzazione con il livello del nero corrispondente a 16 e il livello di picco del bianco corrispondente a 235. Il livello del segnale può occasionalmente superare 235 225 livelli di quantizzazione centrato rispetto alla scala di quantizzazione con il livello zero corrispondente a 128	
Utilizzo delle parole di codice	Le parole di codice corrispondenti ai valori numerici 0 e 255 sono usati esclusivamente per la sincronizzazione. I valori da 1 a 254 sono disponibili per il video	

Tab. 1 - Valori dei parametri di codifica per i membri 4:2:2, 13,5 MHz della famiglia. Sono utilizzati sia nel caso di televisione digitale con formato d'immagine 4:3, che nel caso di utilizzo con formato di visualizzazione 16:9.

5. I PROBLEMI DI CONVERSIONE

Inizialmente il video digitale veniva acquisito e utilizzato soprattutto nel formato 4:2:2, con formato di immagine 4:3. L'evoluzione delle tecnologie digitali in campo video è stata molto rapida e oggigiorno è possibile produrre e post-produrre a costi limitati e con qualità elevata (in molti casi, sia per effetti speciali che per correzioni cromatiche è necessario operare in 4:4:4) ed è sempre più diffuso il formato d'immagine 16:9 (supportato dal DVB, dal DVD e fruibile mediante i nuovi televisori a grande schermo). Diventano sempre più frequenti le conversioni dei vari membri della famiglia, fra di loro e verso i formati di ripresa e di editing non lineare semiprofessionale e consumer.

Limitandosi ai soli ambiti di applicazione della Rac. ITU-R BT.601, si può comunque far cenno ai seguenti problemi che, solo se noti e tenuti in considerazione, consentono di sfruttare al meglio le caratteristiche di qualità dei segnali video digitali:

- La corrispondenza fra i livelli compresi fra nero e bianco è in ambito video, come si è visto precedentemente, è ristretta alle rappresentazioni numeriche fra 16 e 235, ma questa limitazione non è spesso presa in considerazione dai sistemi di grafica elettronica (in genere di derivazione informatica, per i quali al nero corrisponde 0 e al bianco corrisponde 255).

Fig. 6 - Una ripresa può essere effettuata con formato d'immagine 16:9 (a) oppure 4:3 (c), in entrambi i casi può essere rappresentata da 720 campioni in orizzontale e 576 righe. Per visualizzare correttamente le immagini occorre avere diplay opportuni e adottare la corretta deflessione (nel caso di display analogico) o interpolazione (nel caso di display digitale). L'immagine 16:9 riprodotta ha ovviamente una definizione orizzontale inferiore.
Le figure (b) e (d) rappresentano le immagini nel caso in cui si utilizzi un display a pixel quadrato, senza interpolazione, per visualizzare rispettivamente (a) e (c): vi è una deformazione spaziale dell'immagine particolarmente evidente nel caso del formato 16:9 (b), ma presente anche nell'altro caso (d) poiché 720 e 576 non sono in rapporto 4:3.

- L'insieme dei colori rappresentabili nella forma Y, C_R e C_B è più ampio di quello rappresentabile con R, G e B: ciò può avere delle implicazioni sulla generazione ed elaborazione di immagini in ambito digitale. E' opportuno limitare i segnali Y, C_R e C_B, sacrificando eventualmente la saturazione, prima che, convertiti in R, G e B, possano dare origine a degradamenti non desiderati e percettibili su luminanza e tonalità di colore.

- Nel caso di postproduzioni complesse le operazioni di interpolazione e decimazione necessarie per passare da 4:2:2 a 4:4:4 possono dare origine a degradamenti percepibili. D'altro canto è possibile che si verifichino conversioni multiple di questo tipo, poiché i sistemi di videoregistrazione professionali sono in genere di tipo 4:2:2 e quelli di derivazione consumer riducono ulteriormente la banda dei segnali differenza-colore.

- Il rapporto fra il numero di campioni attivi per riga (720) e quello delle righe attive (576) non corrisponde a 4:3, ovvero 1,33, bensì a 1,25: si dice che il pixel non è "quadrato", a differenza di quanto avviene nel caso di sistemi fotografici o di grafica elettronica. Le conseguenze di questo fatto sono che un'immagine 4:3 video può risultare lievemente distorta sullo schermo di un computer, come percepibile confrontando le figure 6c e 6d, oppure che immagini create in ambito informatico appaiano deformate sullo schermo televisivo. In particolare le forme geometriche (cerchi e quadrati) possono apparire visibilmente deformate (ellissi e rettangoli). Per ovviare a questi inconvenienti, sia in fase di produzione che di visualizzazione, i programmi grafici (figura 7) oggigiorno prevedono che si possano impostare automaticamente, grazie al menu utilizzato nella definizione delle dimensioni dell'immagine, il numero di pixel e di righe corretti per il sistema video su cui verrà visualizzata l'immagine (ad esempio prevedendo le voci predefinite PAL e PAL *widescreen*) e viene attivata un opzione per la correzione della proporzione dei pixel in modo che la visualizzazione sullo schermo del computer (a pixel quadrati) sia conforme a quella finale, sul display del televisore.

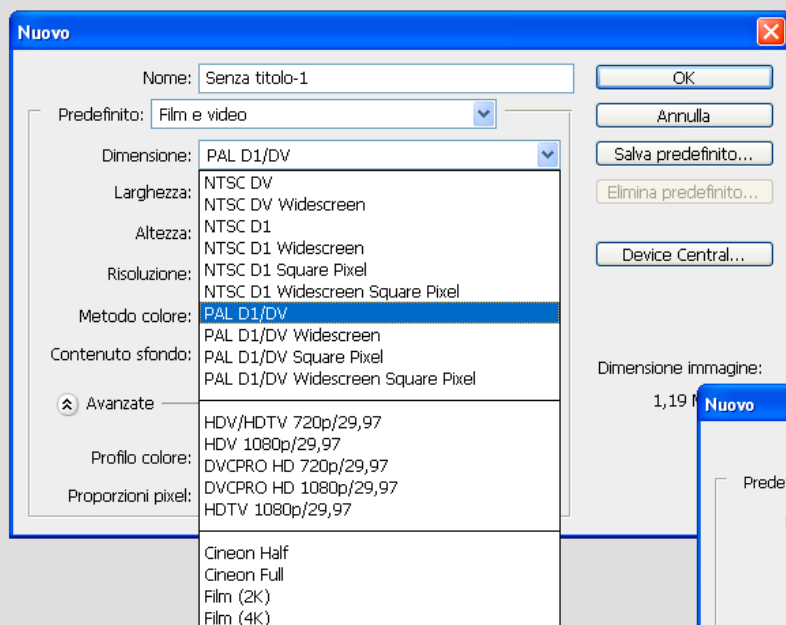



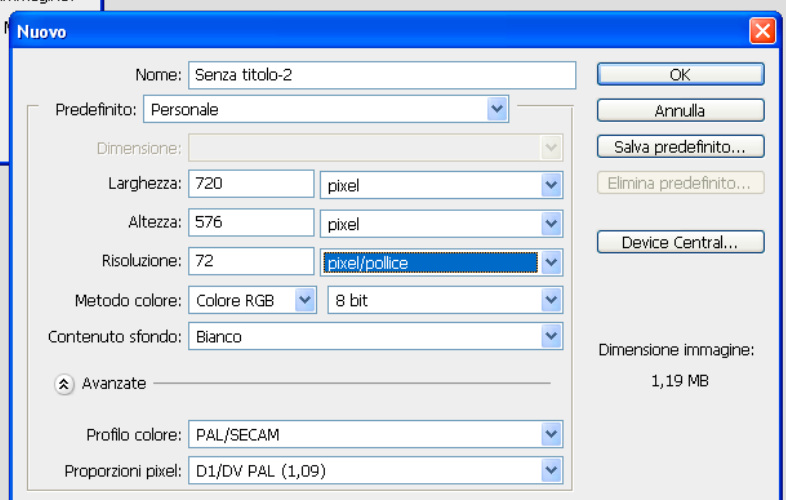
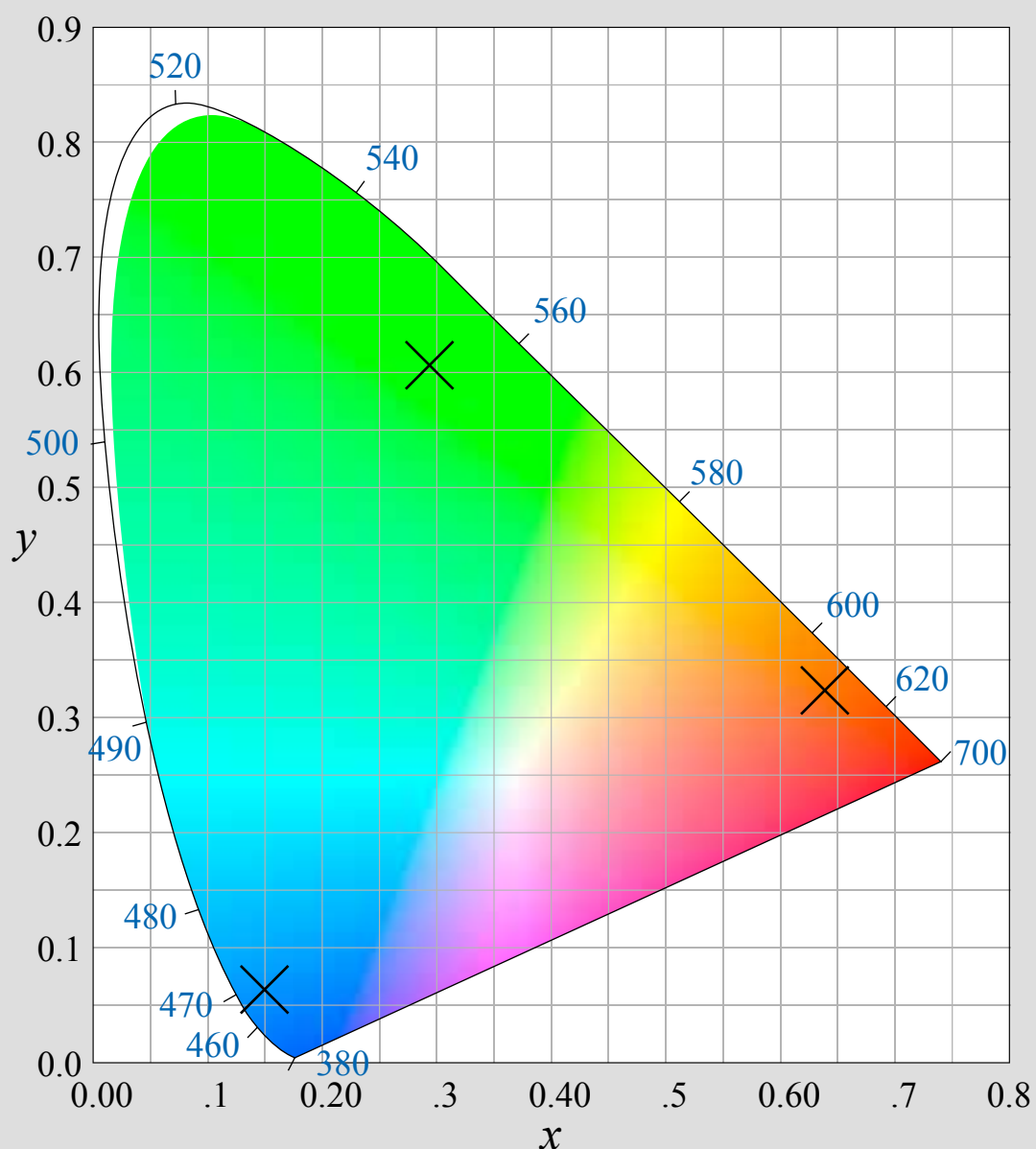
Fig. 7 - Esempio di menu di programma per realizzazione di immagini grafiche (Photoshop) per la scelta del formato di immagine.

Colore primario	ITU BT.601 [2]				ITU-R BT.1361 [3] ITU-R BT.709 [4]	
	625 righe		525 righe			
	x	*y*	*x*	*y*	*x*	*y*
R - Rosso	0,640	0,330	0,630	0,340	0,640	0,330
G - Verde	0,290	0,600	0,310	0,595	0,300	0,600
B - Blu	0,150	0,060	01,55	0,070	0,150	0,060

Fig. 8 - In altro il diagramma CIE 1931 dello spazio cromatico. Le coordinate colorimetriche CIE 1931 specificate dalla Rac. 601 per i colori primari, riportati nella tabella, sono le stesse previste per sistemi analogici a 625 e 525 righe, ma la recente versione della Rac. consente l'uso della colorimetria e dei parametri per la matrice colorimetrica definita dalla Rac. ITU-R BT.1361, al fine di permettere la compatibilità diretta con i sistemi HDTV.

3 Formato d'immagine 16:9 Problemi di conversione

Testo e figure tratte dall'articolo di Massimo Visca, Elettronica e Telecomunicazioni, Dicembre 2009.

Premessa

Il formato 16:9 (panoramico o *widescreen*) è destinato a sostituire nel periodo medio-breve, sia in produzione sia in trasmissione, il tradizionale formato 4:3.

Tra i motivi che rendono ormai inevitabile tale migrazione si possono ricordare:

- la sempre maggior penetrazione sul mercato consumer di schermi in formato 16:9. Entro pochi anni in Italia saranno disponibili a casa degli utenti più display televisivi 16:9 che 4:3, e l'utilizzo di *widescreen* sta diventando comune anche per applicazioni informatiche;
- la possibilità offerta dalla piattaforma digitale terrestre (DVB-T), destinata a sostituire la tradizionale trasmissione analogica entro il 2012, di differenziare l'offerta e fornire all'utente già da oggi il segnale in formato 16:9;
- l'uso ormai consolidato del formato 16:9 da parte dei produttori e broadcaster europei ed extra-europei, e la conseguente creazione di un mercato basato su tale formato;
- l'impiego di formati in alta definizione, disponibili solo in formato 16:9, in tutti i grandi eventi;
- la maggior vicinanza del formato 16:9, rispetto al 4:3, ai rapporti di forma utilizzati per i materiali originati in pellicola o nel cinema digitale, con conseguenti vantaggi nelle operazioni di conversione da pellicola a segnale televisivo.

Questo articolo si pone i seguenti obiettivi:

- fornire alcuni elementi utili ad approfondire la conoscenza dei formati 4:3 e 16:9;
- analizzare i problemi legati alla conversione tra i due formati;
- fornire una terminologia per quanto possibile univoca per identificare le diverse modalità di conversione e visualizzazione;
- analizzare in sintesi il concetto di ripresa protetta;
- riassumere i punti salienti relativi al rispetto delle *Safe Areas* come definite dalla raccomandazione EBU R-95 [6];
- riassumere le modalità adottate dalla Rai per la trasmissione in *simulcast* del tradizionale segnale analogico e del segnale digitale terrestre.

L'articolo introduce solo gli elementi indispensabili per la descrizione del formato 16:9 e delle problematiche di produzione associate, senza voler ambire ad una descrizione esaustiva del tema dal punto di vista tecnico.

1. Segnale televisivo con rapporto di forma 4:3

Il rapporto di forma 4:3 fu utilizzato fin dagli anni 50, contestualmente alla nascita del sistema televisivo come ancora oggi utilizzato, in quanto garantiva un ottimo compromesso tra la tecnologia dei tubi a raggi catodici dell'epoca e la necessità di avere un rapporto di forma in grado di soddisfare le esigenze di composizione dei contenuti dell'immagine.

Ancora oggi, il rapporto di forma 4:3 viene normalmente utilizzato per la produzione e la trasmissione del segnale analogico su tutte le reti Rai e costituisce la quasi totalità del materiale di archivio.

In figura 1 è rappresentato il classico segnale 4:3, dove ovviamente il valore 4:3 si riferisce al rapporto tra la dimensione orizzontale e quella verticale dell'immagine[Nota 1].

Tutta la catena di generazione (telecamera), manipolazione (mixer, videoregistratori, ...) e visualizzazione (monitor o display) del segnale è progettata in modo da mantenere inalterato il rapporto di forma e quindi la geometria dell'immagine.

Nota 1 - I termini "immagine", "quadro video", "trama video" e, in inglese, "*image*", "*picture*", "*frame*" si possono considerare equivalenti. Gli standard europei prevedono che il segnale video sia costituito da 25 *frame* al secondo.

I segnali televisivi utilizzati in Rai utilizzano il rapporto di forma 4:3 sebbene, come meglio specificato in seguito, tanto i segnali analogici (PAL o componenti) quanto quelli digitali (componenti su interfaccia SDI) siano in grado di gestire anche il segnale con rapporto di forma 16:9.

La normativa internazionale [2]prevede per i segnali digitali con rapporto di forma 4:3 una risoluzione pari a 720 pixel in orizzontale e, nel caso di diffusione televisiva utilizzando il sistema MPEG-2 [7], normalmente le immagini codificate sono costituite da 702 pixel in orizzontale e 576 linee in verticale.

2. Segnale televisivo con rapporto di forma 16:9

Il segnale 16:9 è stato introdotto nella produzione televisiva negli anni '80 con lo scopo di aumentare l'angolo di visione dell'osservatore e garantirgli una visione più panoramica ed "immersiva". In ambito cinematografico si utilizzano formati panoramici fin dagli anni 50[Nota 2].

In Europa il formato 16:9 si è andato diffondendo con discreto successo, mentre in Italia è rimasto ai margini della produzione.

Nota 2 - Curiosità: su richiesta delle case di produzione di Hollywood, è stato recentemente introdotto sul mercato un display consumer in formato 21:9.

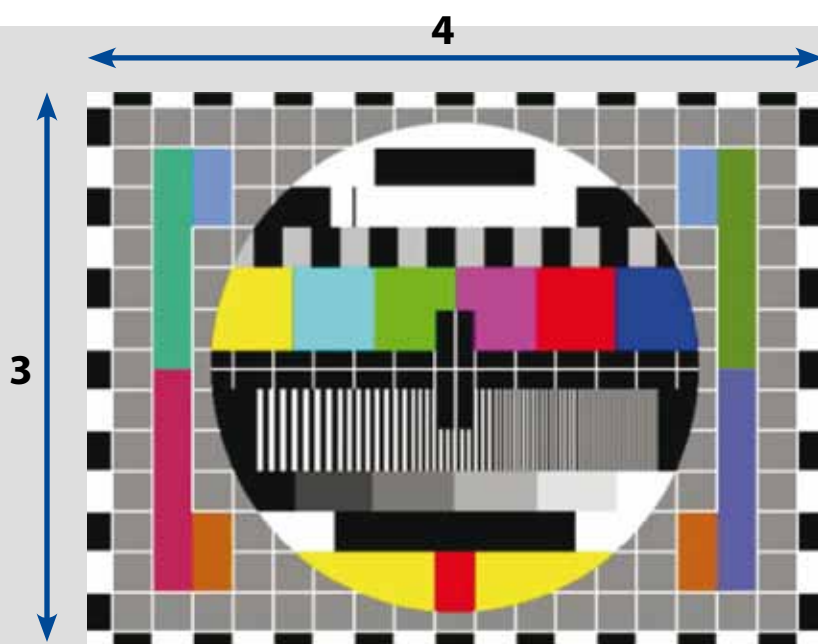



Fig. 1 - Segnale con rapporto di forma 4:3.

In figura 2 è rappresentato un segnale con rapporto di forma 16:9, dove il valore 16:9 si riferisce al rapporto tra la dimensione orizzontale e quella verticale dell'immagine.

Si noti che il rapporto 4:3 può anche essere scritto come 12:9, quindi, a parità di altezza dello schermo, un segnale in formato 16:9 risulta più ampio di un fattore 1,33 rispetto al formato 4:3 (12:9).

Ovviamente, per generare un'immagine 16:9 è necessario un mezzo di ripresa appositamente progettato per gestire tale formato.

Le telecamere moderne[Nota 3] sono di solito in grado, variandone semplicemente la configurazione, di generare segnali 4:3 oppure 16:9.

I display, per garantire il massimo delle prestazioni, devono ovviamente avere un rapporto di forma coerente con quello dell'immagine da visualizzare, sebbene sia possibile, come ampiamente descritto in seguito, utilizzare display 16:9 per visualizzare immagini 4:3, e viceversa.

Per quanto riguarda il resto della catena produttiva, da un punto di vista tecnico, i segnali 4:3 e 16:9, pur essendo diversi tra di loro in termini di contenuto, in quanto l'immagine ripresa e visualizzata è diversa, sono identici dal punto di vista dei parametri tecnici e vengono detti in gergo "elettricamente indistinguibili". Ciò significa che un apparato (mixer, videoregistratore etc) sarà in grado di trattare indifferentemente un segnale 4:3 o 16:9, senza "rendersi conto" della differenza, gestendo il segnale 16:9 nella modalità cosiddetta "4:3 anamorfico", come descritto al successivo paragrafo 4.2.

La normativa internazionale [2] prevede per il rapporto 16:9 lo stesso numero di pixel e linee utilizzato per il formato 4:3, il che implica una minor risoluzione orizzontale dell'immagine 16:9 rispetto a quella 4:3[Nota 4].

3. Segnali televisivi in alta definizione

I segnali citati nei precedenti paragrafi 1 e 2 sono segnali a definizione convenzionale (SDTV - *Standard Definition Television*).

I segnali SDTV si stanno avviando verso una rapida obsolescenza e sono ormai sostituiti da segnali in alta definizione (HDTV - *High Definition Television*).

Nota 3 - il camcorder BetacamSP NON è in grado di generare segnali 16:9. Tale sistema di videoregistrazione professionale è stato ampiamente utilizzato dai broadcaster e molto del materiale di archivio è basato su tale formato.

Nota 4 - La precendente versione della raccomandazione prevedeva anche un segnale 16:9 con 960 pixel, ma tale formato non è mai stato effettivamente utilizzato ed è stato di conseguenza eliminato dalla normativa.

Fig. 2 - Segnale con rapporto di forma 16:9.

Fig. 3 - Segnale in formato 4:3 visualizzato su display 4:3.

Esistono diversi formati HDTV ma TUTTI i segnali in alta definizione utilizzano il rapporto di forma 16:9. Ne consegue che volendo convertire un segnale HDTV in un segnale SDTV (operazione detta *down conversion*) senza modificare il contenuto dell'immagine, si otterrà un segnale SDTV anch'esso 16:9.

4. Conversioni di formato da 4:3 a 16:9 e viceversa

La coesistenza di segnali e display con rapporti di forma diversi genera un'ampia casistica di situazioni in cui occorre convertire un segnale dal rapporto di forma 4:3 al 16:9, o viceversa.

Tutte le conversioni descritte in questo paragrafo possono potenzialmente avvenire o nella sola fase di visualizzazione, per adattare il rapporto di forma del segnale a quello del display, o allo scopo di generare una nuova versione del segnale in formato diverso e destinato ad essere memorizzato, elaborato o trasmesso. E'evidente come questo secondo caso sia più critico e richieda maggior attenzione alla qualità finale del processo.

Nel seguito sono elencati i casi principali.

4.1 Gestione segnale 4:3

Segnale prodotto in formato 4:3 e visualizzato su display 4:3

Se il formato del display è coerente con quello di produzione non si verifica ovviamente alcun problema. La geometria dell'immagine è mantenuta e l'immagine copre completamente la superficie del display. Si tratta ovviamente del formato televisivo da sempre in uso in Italia.

Segnale prodotto in formato 4:3 e visualizzato su display 16:9

La conversione di un formato 4:3 in 16:9 può essere effettuata secondo diverse modalità, riassunte nel seguito.

Modalità *Pillar Box*

In questo caso si mantiene l'immagine in formato 4:3 e si utilizza solo una porzione dello schermo per

Fig. 4 - Modalità Pillar Box o Side Panels.

Fig. 5 - Modalità *Widescreen*.

visualizzarla, riempiendo con due bande nere (o grigie) laterali la porzione di schermo inutilizzata. Non si verifica alcuna perdita di informazione, in quanto l'immagine è adattata perfettamente sulla dimensione verticale dello schermo e non si verifica alcuna distorsione geometrica. La resa estetica è però penalizzata dalla presenza delle bande laterali.

Questa modalità è talora indicata anche con i termini *Side Panels* o *Adaptive*.

Modalità *Widescreen*

In questo caso si mantiene la dimensione verticale dell'immagine ma la si deforma in senso orizzontale per riempire tutto lo schermo. Non si verifica alcuna perdita di informazione ma si introduce una notevole distorsione geometrica sull'immagine, che risulta "stirata" in orizzontale. Sebbene questa modalità risulti penalizzante per la qualità dell'immagine, sembra sia quella più utilizzata dai telespettatori per adattare l'immagine 4:3 sul display 16:9.

Questa modalità è talora indicata anche con i termini 16:9 anamorfico, *Adaptive* o *Squeeze*.

Modalità *Zoom*

In questo caso si effettua un ingrandimento dell'immagine per riempire tutto lo schermo senza deformarne la geometria. Si verifica perdita di informazione in funzione dell'entità dello zoom applicato.

Questa modalità, in ambiente grafico e di post-produzione, viene anche definita *Center Zoom* o *Center Cut*.

Fig. 6 - Modalità Zoom.

Modalità proprietarie

Sono inoltre spesso disponibili sui display soluzioni proprietarie, proposte con nomi diversi dai costruttori.

Alcune di esse applicano distorsioni non lineari sull'immagine, per esempio deformandone solo i lati e mantenendo inalterata la geometria della porzione centrale, dove si presume siano concentrate la parte utile dell'immagine e l'attenzione dell'osservatore.

I risultati non sono sempre positivi, per esempio quando si deforma in modo rilevante un logo grafico posto ai margini dell'immagine o, peggio, si pregiudica una scelta artistica che sfrutta il bordo dell'immagine.

4.2 Gestione del segnale 16:9

Segnale prodotto in formato 16:9 e visualizzato su display 16:9

Non si verifica ovviamente alcun problema, la geometria dell'immagine è mantenuta e l'immagine copre completamente la superficie del display.

Segnale prodotto in formato 16:9 e visualizzato su display 4:3

Nel caso in cui la produzione sia effettuata in formato 16:9, ma sia necessario post-produrre o trasmettere in formato 4:3, il broadcaster deve eseguire presso i propri impianti la conversione necessaria.

Fig. 7 - Riassunto principali conversioni da formato 4:3 a 16:9.

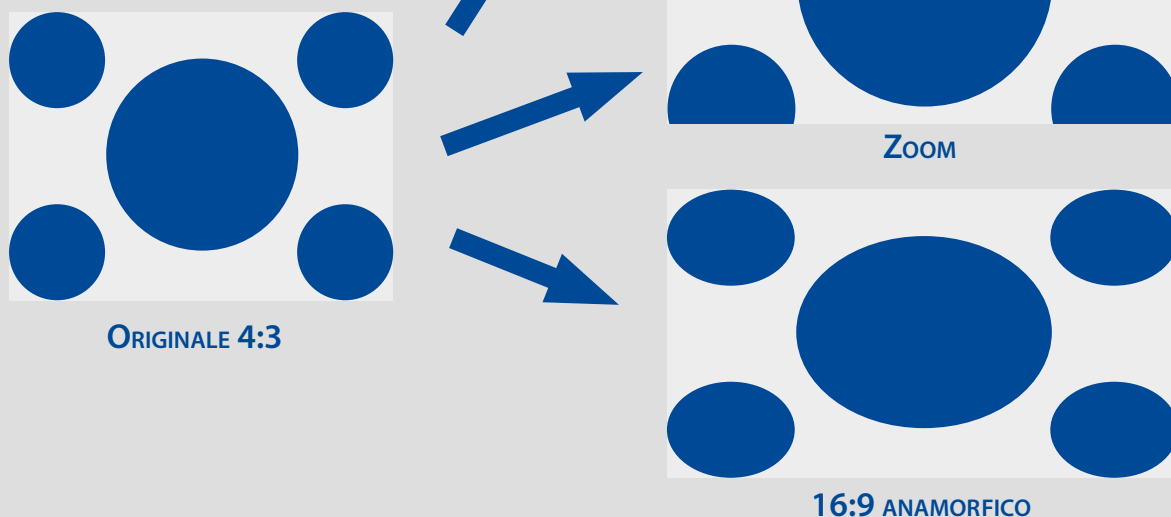



Fig. 8 - Segnale 16:9 visualizzato su display 16:9.

E' questo il caso classico dei prodotti derivati da pellicola quali film e fiction.

Con l'introduzione delle riprese in HDTV, ormai normalmente impiegate per gli eventi sportivi e di intrattenimento e che, come ricordato, utilizzano solo il rapporto di forma 16:9, questo caso d'uso sta diventando sempre più frequente anche nell'ambito della produzione televisiva generalista.

La conversione dal formato 16:9 a 4:3 è quindi un'operazione tipicamente effettuata con apparati professionali (ARC - *Aspect Ratio Converter*) che possono essere integrati in sistemi più complessi (Mixer, Videoregistratori) o essere disponibili come unità hardware singole.

Modalità *Letter Box*

In questo caso l'immagine 16:9 viene trasformata in un formato 4:3 adattando l'immagine sulla dimensione orizzontale e mantenendo tutta l'informazione. La porzione del quadro video "inutilizzata" sopra e sotto l'immagine viene automaticamente riempita con due bande nere.

Non si verifica alcuna perdita di informazione né si verifica alcuna distorsione geometrica sull'immagine. La resa estetica è però penalizzata dalla presenza delle bande nere e dalla perdita di risoluzione verticale (in quanto una parte delle righe televisive vengono utilizzate per descrivere le bande nere). Inoltre, la conversione in Letter Box fornisce un'immagine fortemente penalizzata qualora fruita su un display di dimensioni medio-piccole.

Fig. 9 - Segnale 16:9 visualizzato in formato Letter Box.

Fig. 10 - Segnale 16:9 visualizzato in formato Edge Crop.

Modalità *Edge Crop*

In questo caso si trasforma l'immagine 16:9 in un formato 4:3 mantenendo la dimensione verticale ed eliminando quindi due porzioni di immagine ai lati della stessa.

Si noti, nell'esempio riportato in figura 10, come l'immagine del giocatore in possesso di palla (evidenziato in figura 9 e posto a sinistra dell'immagine), venga tagliata, con evidente impatto sul contenuto e sulla qualità dell'inquadratura visualizzata.

La geometria dell'immagine viene mantenuta.

Questa modalità, in ambiente grafico e di post-produzione, viene anche definita *Center Cut*.

Modalità 4:3 anamorfico

In questo caso il segnale 16:9 viene trattato come se fosse un segnale 4:3[Nota 5]. La visualizzazione viene effettuata su un display formato 4:3 adattando l'immagine sulla dimensione verticale e, di conseguenza, deformandola sulla dimensione orizzontale. Il risultato visivo è un'immagine "compressa" sull'asse orizzontale, quindi con geometria deformata, ma che contiene tutta l'informazione visiva del segnale originale 16:9.

Nota 5 - la modalità "4:3 anamorfico" può essere utilizzata per visualizzare un segnale 16:9 su un display 4:3, o per eseguire determinate lavorazioni utilizzando le dovute cautele, ma non può essere usata per la trasmissione all'utente.

Fig. 11 - Modalità di visualizzazione 4:3 anamorfico; l'immagine è compressa sull'asse orizzontale.

Per quanto riguarda le lavorazioni e manipolazioni del segnale 4:3 anamorfico occorre invece distinguere tra le operazioni che agiscono su tutta l'immagine (es. dissolvenza, tendina lineare) e quelle che modificano sono una parte di essa (es. tendina con forma geometrica, titolazione o inserimento di logo).

Nel primo caso gli apparati sono in grado di gestire il segnale senza complicazioni e, per esempio, il mixer potrà far transitare il segnale 16:9 ed applicare tendine lineari e dissolvenze, il videoregistratore potrà registrare e riprodurre il segnale 16:9[Note 6,7].

Viceversa, nel caso in cui si intarsino forme grafiche (tendine con forme geometriche quali cerchi o quadrati, caratteri di testo, loghi etc) generate secondo proporzioni adatte al formato 4:3, si incorre ovviamente nel problema che visualizzando l'immagine su un display 16:9 le forma grafiche inserite risulteranno distorte ed allargate in orizzontale.

Per evitare questo problema occorre inserire sull'immagine 4:3 anamorfica una forma grafica anch'essa 4:3 anamorfica. Alcuni apparati, per esempio le titolatrici moderne, supportano questa funzionalità.

Nota 6 - i videoregistratori BetacamSP ed IMX gestiscono senza alcun problema il segnale 4:3 anamorfico.

Nota 7 - i sistemi di editing non lineare, per esempio le piattaforme AVID, gestiscono senza alcun problema i formati 4:3 e 16:9.

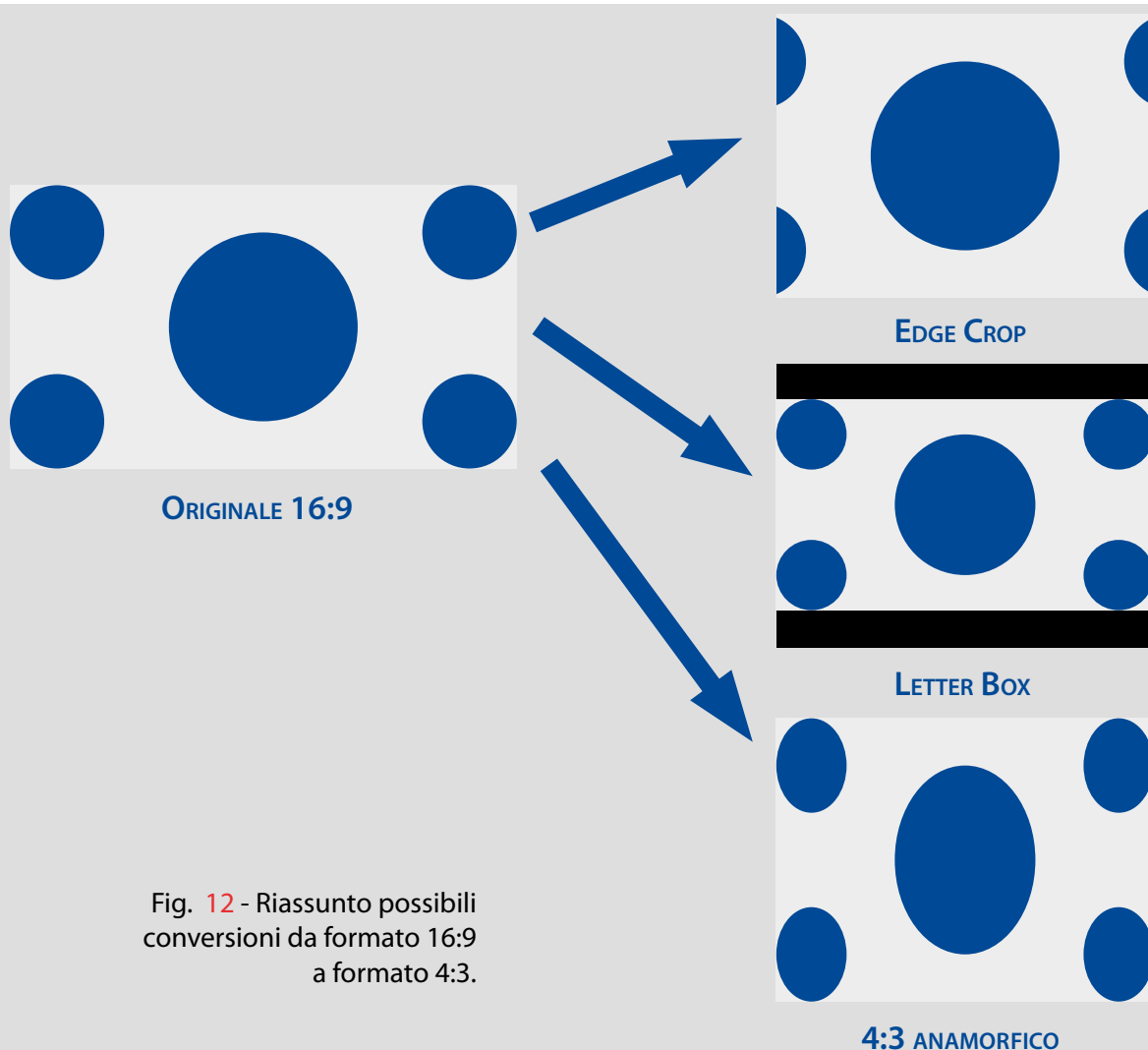



Fig. 12 - Riassunto possibili conversioni da formato 16:9 a formato 4:3.

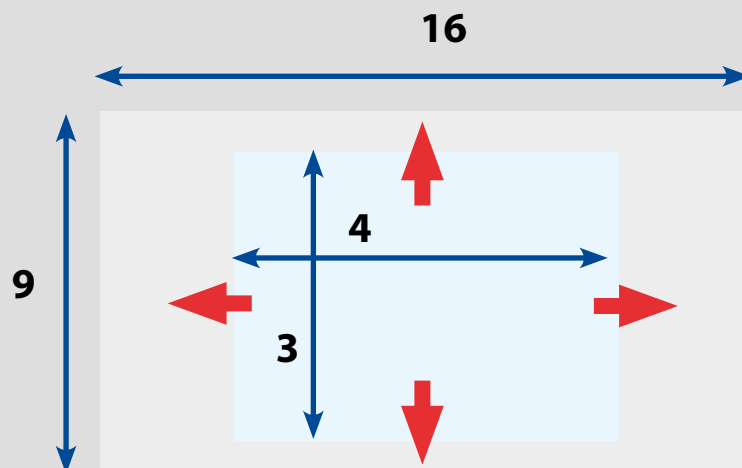



Fig. 13 - Modalità *Pan & Scan*.

Modalità Pan & Scan

In casi particolari si può effettuare la conversione selezionando, scena per scena, una finestra di formato 4:3 all'interno dell'immagine 16:9, in modo da salvaguardare nel miglior modo possibile il contenuto del programma.

Questa tecnica, definita di solito *Pan & Scan* è ovviamente costosa e si utilizza solo in post-produzione per convertire film e fiction, mentre non trova applicazioni nel caso di eventi in diretta.

5. Ripresa protetta o Protected shooting

La conversione di un formato 16:9 in 4:3 effettuata con la modalità *Edge Crop* comporta l'eliminazione irreversibile di una porzione significativa dell'immagine sulla dimensione orizzontale, pari al 25% della superficie del quadro video.

Come evidenziato dal confronto delle figure 8 e 10, l'immagine 4:3 può risultare fortemente penalizzata dal punto di vista del contenuto, qualora il soggetto principale si trovi ai lati dell'inquadratura.

Per evitare questo inconveniente, qualora si riprenda in formato 16:9 sapendo che le immagini dovranno essere utilizzate secondo la modalità di *Edge Crop*, può essere adottata la cosiddetta ripresa protetta, che consiste nel costruire l'inquadratura posizionando gli elementi principali della stessa in un immagine virtuale in formato 4:3 evidenziato all'interno dell'inquadratura 16:9.

La ripresa protetta tutela il formato 4:3 ma, paradossalmente, può risultare penalizzante per i fruitori dell'immagine 16:9, come evidenziato dall'esempio riportato in figura 14.

Fig. 14 - Ripresa protetta; sono evidenziati i margini del formato 4:3.

Con la sempre maggior diffusione dei formati HDTV è quindi probabile che le riprese protette, specie nel caso di grandi eventi prodotti per mercati dove l'alta definizione è ormai affermata, siano sempre meno utilizzate e tollerate.

6. Safe Areas

La raccomandazione [6] prevede (*recommends*) che, in caso di produzione in formato 16:9, l'inquadratura e la grafica rispettino le *Safe Areas* riportate in figura 15 e, in particolare, che:

- il contenuto informativo dell'immagine (*essential action*) sia interamente compreso nel perimetro della *Action Safe Area*;
- tutti i contributi grafici siano compresi nel perimetro della *Graphics Safe Area*;
- il centro del immagine mantenga la stessa posizione durante il processo di manipolazione del segnale, a meno che non vi siano motivi artistici per violare intenzionalmente tale regola.

Le *Safe Areas* hanno lo scopo di salvaguardare i contenuti del immagine e garantire che vengano visualizzati anche su display che eliminano i bordi dell'immagine a causa del processo di overscan [Nota 8].

> Nota 8 - il rispetto delle *Safe Areas* non deve essere in alcun modo confuso con la modalità di ripresa *Protected Shooting* trattata nel precedente Paragrafo 5.

7. Trasmissioni Rai in formato 4:3 e 16:9

Fino a pochi anni or sono, il parco ricevitori domestici era costituito per la quasi totalità da televisori in formato 4:3 e tale fatto ha precluso ogni possibilità per i broadcaster italiani di trasmettere in formato 16:9 e, di conseguenza, anche gli spazi per produrre in formato panoramico sono sempre stati molto limitati.

Questa situazione è cambiata con la graduale penetrazione sul mercato dei display a schermo piatto, in larga misura in formato 16:9, e con l'introduzione della piattaforma di trasmissione digitale terrestre e del relativo Set Top Box necessario alla ricezione.

Il rilevante numero di variabili in gioco genera un elevato numero di possibili casi d'uso, e lo scenario è ulteriormente complicato dal fatto che la migrazione da segnale analogico a digitale avverrà in un arco temporale di alcuni anni.

Al fine di soddisfare al meglio l'utenza e sfruttare le potenzialità offerte dalle nuove tecnologie, i broadcaster hanno sostanzialmente identificato due principali modalità di fruizione dell'immagine televisiva:

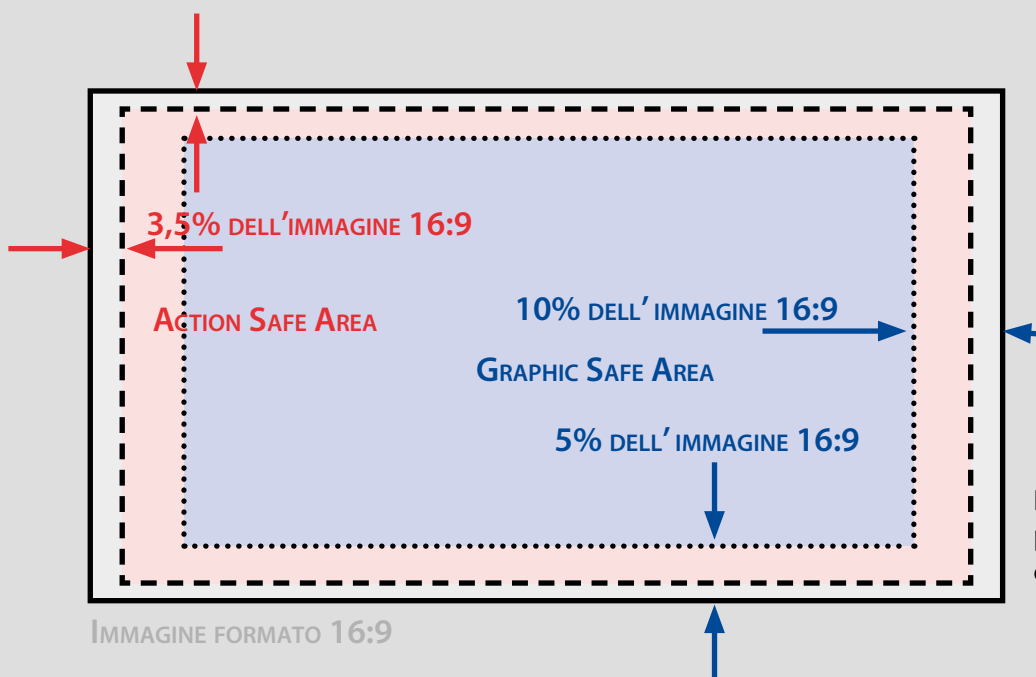



Fig. 15 - Safe Areas per il formato 16:9 come definite in [6].

- la prima, basata sulla ricezione tramite rete analogica e visualizzazione con televisore 4:3, che rappresenta la garanzia per gli utenti ancora legati alla vecchia tecnologia di poter continuare a usufruire del servizio tradizionale fino allo spegnimento completo della rete analogica.

- la seconda, basata sulla ricezione del segnale digitale terrestre associata ad un display 16:9, che offre agli utenti in possesso di display *widescreen* la possibilità di sfruttarne appieno le caratteristiche, e che rappresenta la modalità di fruizione del futuro per tutti gli utenti.

Si rende quindi necessario alimentare le due reti, analogica e digitale Nota 9, con segnali con rapporto di forma diverso, generando i casi di fruizione descritti, in sintesi, nel seguito.

7.1 Segnale prodotto in formato 4:3

Nel caso in cui il programma sia disponibile in formato 4:3 (perché così generato in studio o disponibile su supporto di archivio, inclusi i casi descritti in precedenza in cui il formato 4:3 è stato ottenuto per conversione in *Letter Box* o *Edge Crop*), lo stesso viene trasmesso senza alterarne il rapporto di forma sia sulla tradizionale catena di diffusione analogica sia sulla piattaforma digitale terrestre. Si verificano di conseguenza le casistiche di ricezione descritte nel seguito.

Ricezione da rete analogica

I ricevitori in formato 4:3, tipicamente i tubi a raggi catodici, visualizzano il segnale senza modificarlo e quindi senza problemi.

I ricevitori in formato 16:9, tipicamente i display a schermo piatto (plasma o LCD) o i rari display CRT *widescreen* analogici, visualizzano il segnale secondo una delle modalità descritte in precedenza (figura 7), penalizzando quindi in qualche modo i possessori di display panoramici.

Ricezione da rete digitale terrestre

In questo caso la rete digitale risulta sostanzialmente trasparente e si ricade nel caso citato al precedente paragrafo.

7.2 Segnale prodotto in formato 16:9

Gran parte degli studi Rai e degli *OBVan*, e sicuramente tutti quelli di nuova costruzione, sono in grado di produrre utilizzando il formato 16:9. Già oggi molti grandi eventi sono prodotti utilizzando tale formato che viene utilizzato, secondo diverse modalità, per alimentare sia la rete analogica sia quella digitale.

Ricezione da rete analogica

Trasmettere un segnale 16:9 sulla rete analogica terrestre significherebbe costringere tutti i possessori di televisori 4:3 analogici, che sono ancora la maggioranza, a fruire di un segnale 4:3 anamorfico (figura 11), cioè distorto in modo significativo.

Per salvaguardare la ricezione del parco ricevitori 4:3 la soluzione adottata consiste nel convertire, all'atto della messa in onda, il segnale 16:9 in formato 4:3 con le modalità *Edge Crop* (di solito usata per gli eventi sportivi) o *Letter Box* (di solito usata per fiction e film), ed alimentare con tali segnali la rete analogica Nota 10.

Ricezione da rete digitale terrestre

Il segnale 16:9 viene trasmesso sulla rete digitale terrestre associando allo stesso, in fase di messa in onda, un segnale di servizio che informa il ricevitore circa il formato ricevuto. Il ricevitore (*Set Top Box*), a cui può essere fornito dall'utente, o dal display stesso, il rapporto di forma del display a cui è collegato, fornirà in uscita il formato coerente, cioè 16:9 nel caso di display di tale formato, o *Letter Box* o *Edge Crop*, secondo la scelta operata dall'utente, nel caso di display 4:3.

Nota 9 - la trasmissione dello stesso programma su due piattaforme diverse, in questo caso rete analogica e rete digitale, viene spesso indicata con il termine *simulcast*, contrazione dei termini inglesi *simultaneous broadcasting*.

Nota 10 - la conversione viene effettuata in fase di messa in onda da un opportuno apparato, a cui viene fornito il comando di effettuare o meno la conversione da un automatismo legato alla messa in onda. Poiché l'apparato ha tempi di reazione dell'ordine di qualche quadro video (ogni quadro video o *frame* ha una durata pari a 40 ms), è possibile talora apprezzare che all'atto dell'intervento dell'apparato si verifica per un istante la trasmissione del segnale in 4:3 anamorfico e poi la corretta visualizzazione in *Edge Crop*.

Interfacce video (SDI, SDTI, ASI, HD-SDI, DVI, HDMI)

Testo e figure tratte dall'articolo di Marzio Barbero e Natasha Shpuza, Elettronica e Telecomunicazioni, dicembre 2006.
Il testo è aggiornato sulla base delle attuali versioni degli standard.

1. Premessa

La Rac. ITU-R BT.601 [2], oggetto del secondo capitolo, e la Rac. ITU-R BT.709 [4] definisono i formati rispettivamente per i segnali video digitali a definizione standard e ad alta definizione ed in particolare definiscono i parametri fondamentali per rappresentare le immagini, quali la struttura e la frequenza di campionamento (figura 1), il numero di quadri al secondo e il numero di righe che costituiscono ciascuna immagine. Non si occupano, invece, di definire le modalità e le caratteristiche elettriche necessarie per interfacciare i vari apparati che, grazie alla catena di produzione, e attraverso la catena di trasmissione, consentono la visione delle immagini a casa dell'utente.

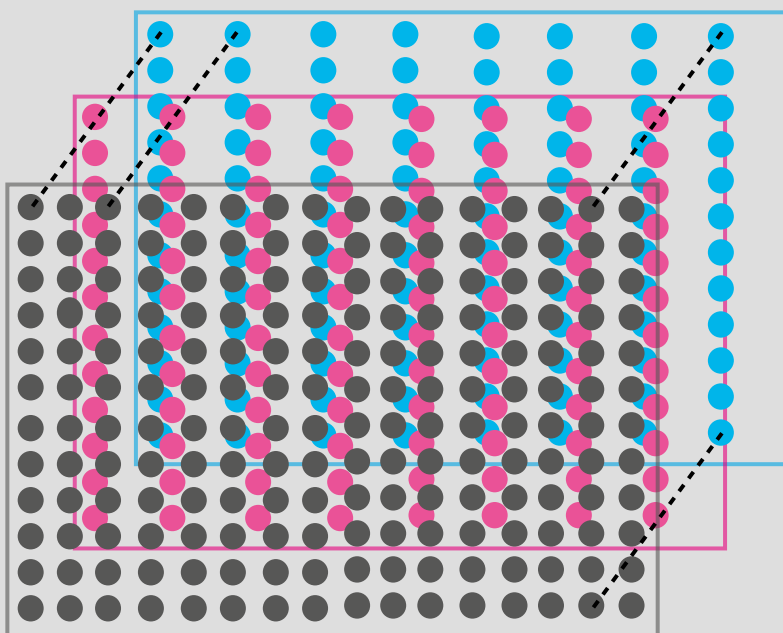



Fig. 1 - In base alla BT.601 il video è campionato secondo la struttura rappresentata a sinistra: ai campioni di luminanza corrisponde una frequenza di 13,5 MHz, a ciascuna delle due componenti di crominanza una frequenza di 6,75 MHz.
I due campioni di crominanza sono coposizionati con i campioni di ordine dispari della luminanza.
I dati (10 bit per campione) relativi ai campioni video sono multiplati, nel seguente ordine: C_B, Y, C_R, Y, C_B, Y, C_R, Y, ...

Fig. 2 - L'interfaccia parallela prevedeva connettori 25-pin D-subminiature e cavi costituiti da 11 doppini (twisted pair), 10 per ciascuno dei bit relativi all'informazione video più uno per il clock. Il cavo poteva raggiungere i 50 m senza equalizzazione e 200 m con appropriata equalizzazione.

Inizialmente, dopo la pubblicazione nel 1982 della Raccomandazione ITU-R BT.601, il problema dell'interfacciamento di tipo digitale si poneva esclusivamente nell'ambiente di produzione, poiché, a quel tempo, il segnale video diffuso era ancora esclusivamente analogico e l'interfacciamento fra gli apparati video domestici era assicurato dalla SCART (§ 5.1).

Le interfacce per la produzione furono definite dagli standard della SMPTE ST 125:1995 (*Television - Component Video Signal 4:4:4 - Bit Parallel Digital Interface*) per i sistemi televisivi a 525 righe e dalle specifiche tecniche EBU Tech 3267 per i sistemi a 625 righe. Entrambe le specifiche confluirono nella Raccomandazione ITU-R BT.656 pubblicata nel 1986.

Nelle prime versioni di tale raccomandazione era prevista una interfaccia parallela (figura 2), infatti i circuiti integrati utilizzati negli anni '80 avevano caratteristiche di velocità tali da rendere difficile la serializzazione del flusso binario. L'attuale revisione, è del 2007 [5] prevede solamente la versione seriale di tale interfaccia, la versione parallela è descritta solamente in appendice, a scopo informativo.

Quando furono per la prima volta interconnessi apparati dotati di queste interfacce, ci si rese conto che la scelta di 13,5 MHz come frequenza di campionamento aveva un inconveniente: la 9ª e la 18ª armonica coincidono esattamente con 121,5 MHz e 243 MHz, canali di emergenza per l'aeronautica. La scelta di questa frequenza era stata il risultato di un compromesso fra diverse proposte, e quindi non era pensabile riconsiderare il valore scelto: per limitare le possibili conseguenze di emissioni spurie a tali frequenze ci si affida a indicazioni rigorose nella progettazione degli apparati e nella cura dei collegamenti.

Successivamente furono definite interfacce adatte per l'interconnessione di apparati anche in alta definizione, sia in ambito professionale (HD-SDI) che domestico (DVI, HDMI).

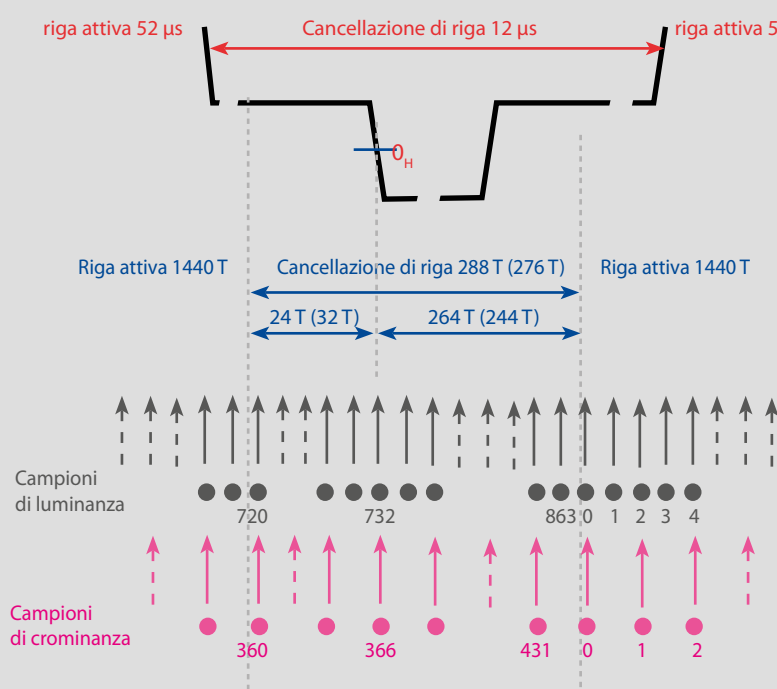



Fig. 3 - Corrispondenza fra i sincronismi digitali e quelli del segnale analogico. Il primo campione della riga attiva digitale è identificato come campione 0, l'ultimo campione è 863 nel caso del formato a 625 righe (857, tra parentesi i valori che differiscono nel formato a 525 righe rispetto a quello a 625 righe. L'intervallo di cancellazione di riga corrisponde a 288 periodi di campionamento (276). Come illustrato nel dettaglio nella figura in alto, il campione corrispondente all'istante O_H (il riferimento per il sincronismo di riga nel segnale analogico) è quello 732 (736) per la luminanza Y e il 366 (368) per le due componenti di crominanza C_B e C_R.
Nella figura in basso è evidenziata la corrispondenza temporale fra i segnali di riferimento digitali SAV e EAV con il sincronismo di riga analogico.

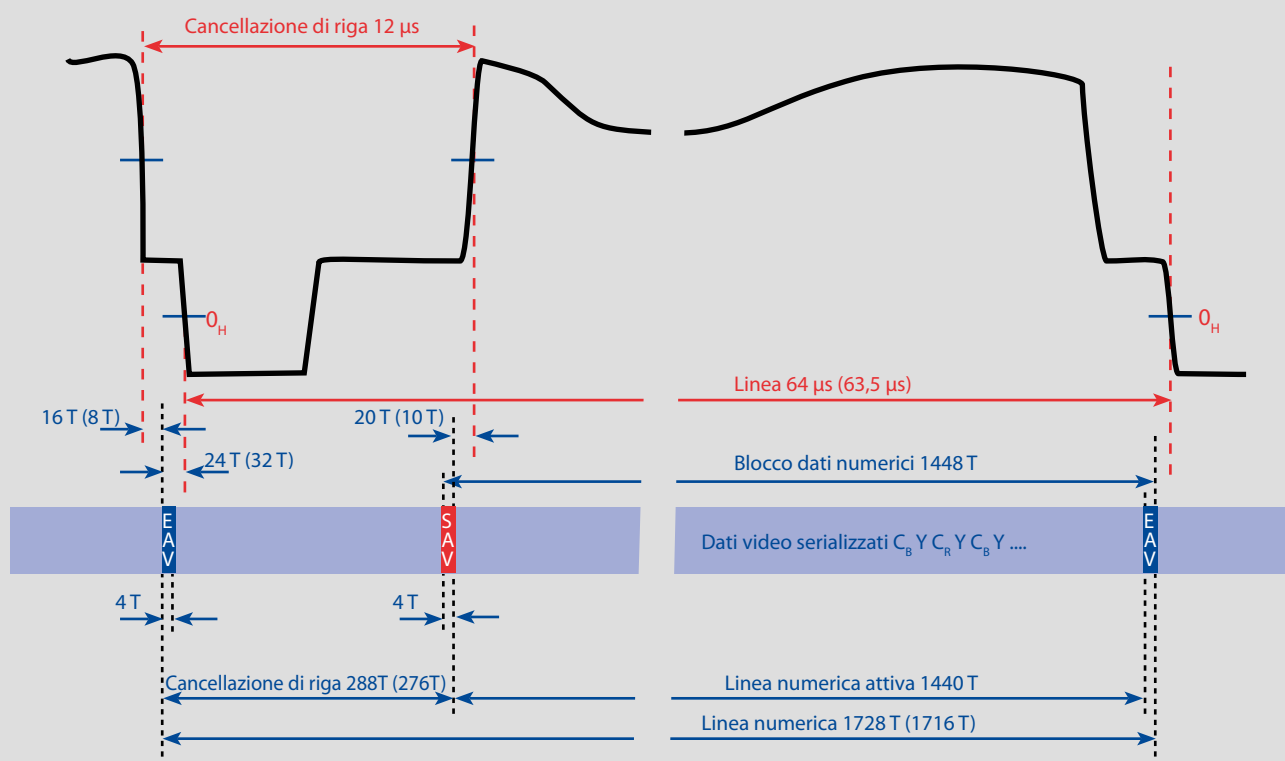



Fig. 4 - SAV e EAV sono i segnali di riferimento, il primo all'inizio di ogni blocco di dati video ed il secondo alla fine del blocco. Ciascun segnale è costituito da quattro parole nel formato come indicato nella prima tabella. Le prime tre parole sono un preambolo fisso costituito da una parola di bit 1 e due parole di bit 0 (queste parole non sono utilizzabili, così come specificato dalla BT.601, per rappresentare informazioni video). Poiché la prima realizzazione dell'interfaccia era a 8 bit, quando è stata estesa a 10 bit, la norma ha indicato come non definiti il valore dei due bit meno significativi, per compatibilità con gli apparati preesistenti, operanti a 8 bit.

Posizione del bit	Prima parola	Seconda parola	Terza parola	Quarta parola
9 (MSB)	1	0	0	1
8	1	0	0	F
7	1	0	0	V
6	1	0	0	H
5	1	0	0	P_3
4	1	0	0	P_2
3	1	0	0	P_1
2	1	0	0	P_0
1	1	0	0	0
0	1	0	0	0

numero di riga video formato 625/50	F	V	H (SAV)	H (EAV)
1-22	0	1	1	0
23-310	0	0	1	0
311-312	0	1	1	0
313-335	1	1	1	0
336-623	1	0	1	0
624-625	1	1	1	0

Per la quarta parola il significato dei bit è quello schematizzato nella seconda tabella: F=0 durante il semiquadro 1 (dalla riga 1) e 1 durante il semiquadro 2 (dalla riga 313, si riportano i dati validi per il formato a 625 righe); V=1 durante l'intervallo di cancellazione di quadro (dalla riga 624 alla riga 23 e dalla riga 311 alla riga 336) mentre è 0 durante le righe video attive; H=0 identifica SAV e H=1 identifica EAV. P_0, P_1, P_2 e P_3 sono bit di protezione, calcolati in base al valore di F, V e H secondo la tabella in basso (il codice consente la correzione degli errori singoli e la rivelazione di quelli pari).

F	V	H	P_3	P_2	P_1	P_0
0	0	0	0	0	0	0
0	0	1	1	1	0	1
0	1	0	1	0	1	1
0	1	1	0	1	1	0
1	0	0	0	1	1	1
1	0	1	1	0	1	0
1	1	0	1	1	0	0
1	1	1	0	0	0	1

2. Raccomandazione 656

2.1 Sincronismi analogici e sincronismi digitali

Per consentire la conversione del segnale video da analogico a digitale e viceversa occorre stabilire univocamente la corrispondenza tra i sincronismi analogici (di quadro, semiquadro e riga) e le informazioni digitali (figura 3).

I sincronismi di riga e di quadro sono sostituiti con due sequenze di 4 parole di 10 bit: SAV e EAV (figura 4).

Il tempo relativo ai restanti campioni della cancellazione di riga (cioè 282 parole nel caso del formato 625/50) è utilizzabile per dati aggiuntivi (*ancillary data*).

Fig. 5 - connettore BNC per cavo coassiale. Deve avere caratteristiche conformi allo standard IEC 61169-8 (2007-2).

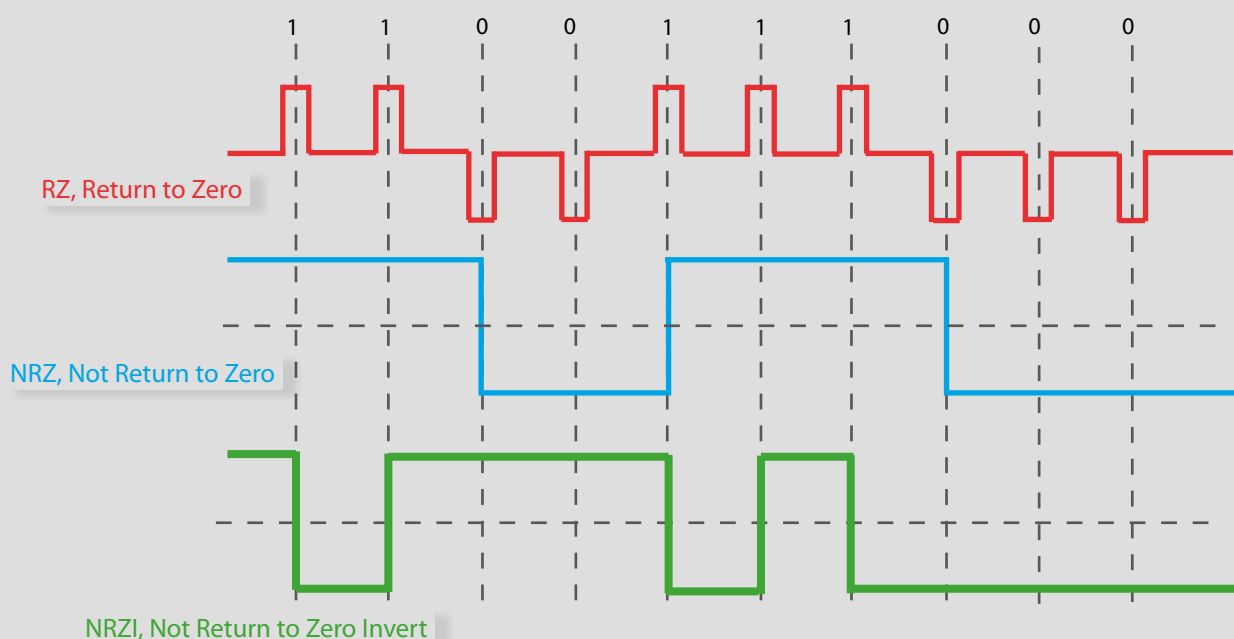



Fig. 6 - Uno dei primi codici di canale utilizzato, in particolare anche per la registrazione dei dati sui primi floppy disk in campo informatico, è stato il codice RZ (*Return to Zero*). In tale applicazione, la corrente di registrazione assume uno stato zero in assenza di bit di informazione, mentre fluisce in una direzione nel caso di bit 1 e nella direzione opposta nel caso di bit 0. Un segnale di questo tipo ha ottime caratteristiche dal punto di vista della rigenerazione del clock, ma non garantisce un buon sfruttamento della banda e ha scarse prestazioni dal punto del rapporto segnale/rumore (la soglia di decisione non è a meta tra i due valori estremi).

Per risolvere questi problemi si può utilizzare il codice NRZ (*Non-Return to Zero*), in questo caso la corrente fluisce in una direzione oppure nell'altra a seconda del valore del bit. I limiti di questo codice sono dovuti al ridotto numero di transizioni (difficile estrazione del clock) e al fatto che ogni qual volta il bit passa da 0 a 1 o da 1 a 0 viene cambiata la polarità del segnale: se si perde un impulso risultano sbagliati non solo il bit immediatamente successivo, ma tutti i bit da tal punto in avanti risultano invertiti (propagazione dell'errore).

Il codice NRZI (*Non-Return to Zero Invert*), risolve il problema della propagazione degli errori assegnando l'evento inversione di polarità alla presenza del simbolo 1, mentre al simbolo 0 corrisponde la mancanza di inversione di polarità. Anche in questo caso, però, il numero di transizioni può essere talmente basso (nel caso di lunghe sequenze di 0) da rendere difficile il recupero del clock.

2.2 Interfaccia seriale (SDI)

L'interfaccia di tipo parallelo, inizialmente prevista da questa raccomandazione, non era praticamente utilizzabile, per realizzare studi televisivi complessi, e quindi l'obiettivo era la definizione di una interfaccia di tipo seriale, che potesse sfruttare le infrastrutture già presenti, basate su cavi coassiali con impedenza a 75 Ω e connettori di tipo BNC (figura 5).

Inizialmente fu sperimentata un'interfaccia seriale in cui la codifica di canale era basata su un codice a gruppo 8/9 (supportava quindi solo 8 bit per campione video), successivamente fu proposto e adottato un sistema basato sull'uso del codice NRZI (figura 6) e di *scrambler* (figura 7). Lo *scrambler* opera sui dati codificati a 10 bit e fa si che i dati assumano una statistica pseudocasuale, caratterizzata da un elevato numero di transizioni: in questo modo è possibile recuperare in ricezione il sincronismo di clock, senza aggiunta di ridondanza.

Fig. 7 - La codifica NRZI più scrambler è utilizzato nell'interfaccia seriale SDI.
La codifica NRZI ha dei vantaggi rispetto alle altre soluzioni, compresi i codici a gruppo, per quanto riguarda l'immunità al jitter. D'altro canto lo spettro è più ampio e non ha un'elevata componente spettrale in corrispondenza con la frequenza di clock. Lunghe sequenze di simboli 0 e 1 nel caso dello NRZ e di 0 nel caso di NRZI non presentano commutazioni: può pertanto risultare molto difficile il recupero della frequenza di clock.
La soluzione è la somma logica (XOR) di una sequenza pseudocasuale ai dati prima della codifica di canale. Una sequenza pseudocasuale è tale se i simboli 0 e 1 sono equiprobabili e hanno statistica simile a quella ottenibile a partire da un generatore di rumore.
E' possibile ottenere una sequenza con statistica pseudocasuale utilizzando un circuito semplice basato su polinomi generatori, gli stessi utilizzati per i codici ciclici CRC. Una volta sommata la sequenza pseudocasuale i dati risultanti hanno una statistica anch'essa di tipo causale, quindi il flusso è caratterizzato da un elevato numero di transizioni e può quindi essere utilizzata una codifica tipo NRZI. L'operazione di somma della sequenza pseudocasuale è spesso indicata come scrambling (il verbo inglese to scramble corrisponde all'italiano "strapazzare").
In ricezione viene sommata la stessa sequenza pseudocasuale (XOR) ottenendo il flusso di dati originali.
Lo scrambler basato su un polinomio di grado 9 è applicato ai dati NRZ nell'interfaccia seriale per il video digitale SDI, seguito da un secondo scrambler di grado 1 per ottenere la sequenza NRZI priva di polarità.

L'interfaccia di tipo seriale è nota comunemente come SDI (*Serial Digital Interface*) e supporta la codifica a 10 bit del formato 4:2:2 con formato d'immagine a 720 pixel attivi per riga (270 MHz) e con formato 16/9 a 960 pixel (360 MHz). Per trasferire un segnale progressivo sono necessari due connessioni BT.656.

Il segnale video seriale, oltre che su cavo coassiale, può essere trasportato su fibra ottica [8].

La norma SMPTE ST 259:2008 (Television – *SDTI Digital Signal/Data - Serial Digital Interface*) prevede anche l'uso per il trasporto di NTSC composito campionato $4f_{sc}$ a 143 Mbit/s, mentre nel caso di PAL composito campionato a quattro volte la sottoportante di colore opera a 177 Mbit/s. La norma ST 259:2008 supporta fino ad otto canali audio digitali AES/EBU.

3. Interfacce seriali per il trasporto del video compresso

3.1 SDTI

Successivamente furono sviluppati sistemi di compressione dell'informazione video, codec e videoregistratori, e ci si pose il problema di interfacciare apparati utilizzanti lo stesso sistema di compressione e formato dati. Infatti l'uso di una interfaccia SDI presuppone che i dati video siano nel formato non compresso BT.601. Il suo utilizzo con informazioni video compresse implica quindi operazioni di co-decodifica in cascata, al solo scopo di interfacciare apparati, anche nel caso questi adottino lo stesso sistema di compressione e formato dati. Poiché i sistemi di compressione utilizzati generalmente non sono trasparenti, comportano una perdita di informazione e un degradamento della qualità video, non giustificato dalla sola esigenza di interfacciamento e di riversamento dei segnali.

Per ovviare a tale limitazione della SDI, la SMPTE ST 305:2005 (*Television – Serial Data Transport Interface (SDTI)*) specifica la SDTI che utilizza l'infrastruttura SDI in modo compatibile sia dal punto di vista elettrico che della temporizzazione e organizza i dati video compressi tra i sincronismi di riferimento SAV ed EAV (figura 8).

I formati video compressi richiedono una velocità di trasferimento inferiore a quella necessaria per il

Fig. 8 - La struttura dati SDI prevede per il video (carico utile) 1440 parole da 10 bit fino ad una capacità massima di 200 Mbit/s nel caso di bit-rate complessivo di 270 Mbit/s e 1920 parole, consentendo fino a 270 Mbit/s per il payload, nel caso di bit-rate complessivo di 360 Mbit/s.
Nel caso di utilizzazione per la SDTI, una intestazione (header) di 53 parole contiene gli indirizzi di sorgente e destinazione e la formattazione del carico utile, mentre il carico utile è organizzato in pacchetti dati denominati item (ne sono previsti quattro tipi: sistema, immagine, audio e ausiliari) ognuno dei quali può contentere fino a 255 elementi. Sono previsti diversi modi di trasferimento, non tutti necessariamente supportati dagli apparati: isocrono (i pacchetti sono agganciati alla struttura di trama SDTI), asincrono, a bassa latenza (struttura organizzata in sottopacchetti per consentire ritardi molto ridotti.

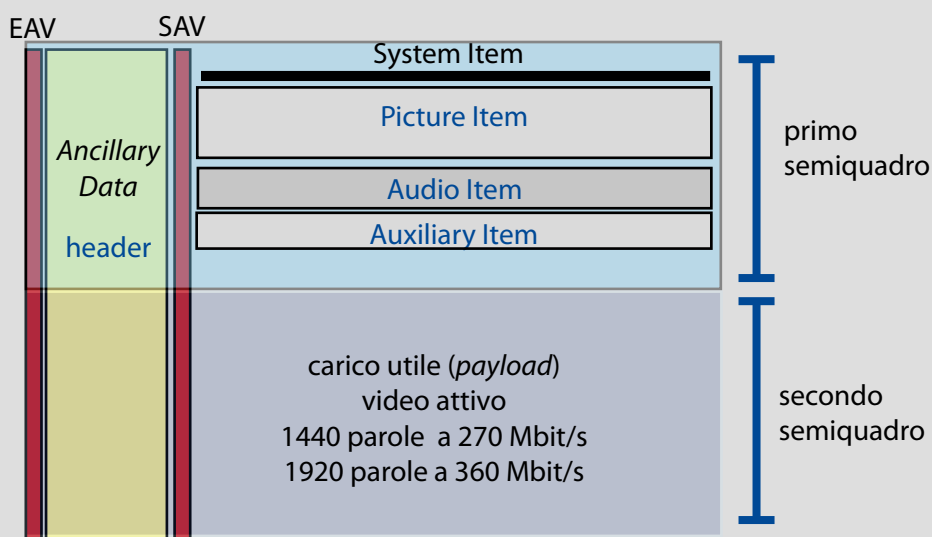
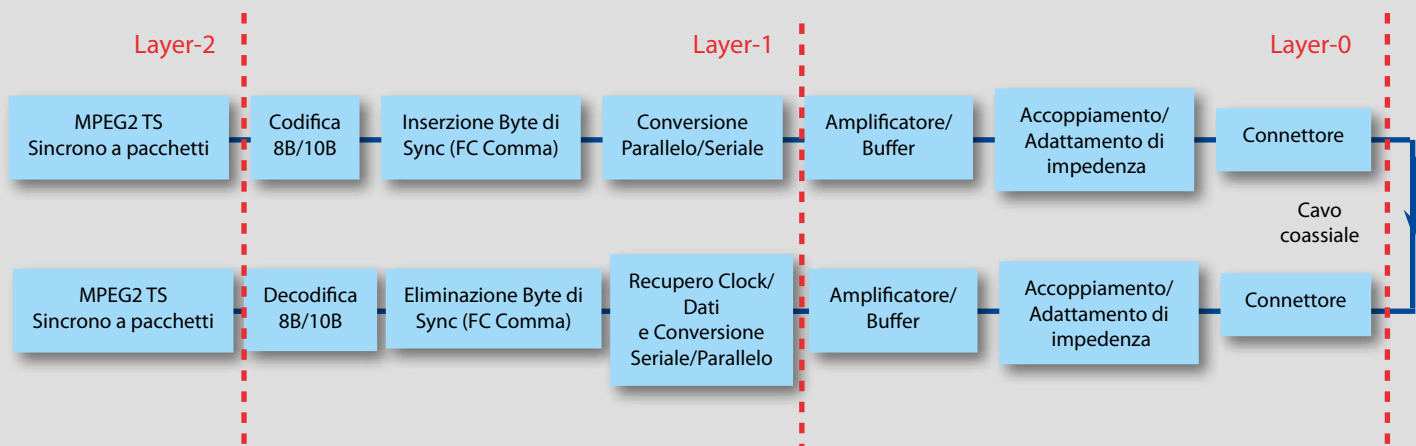
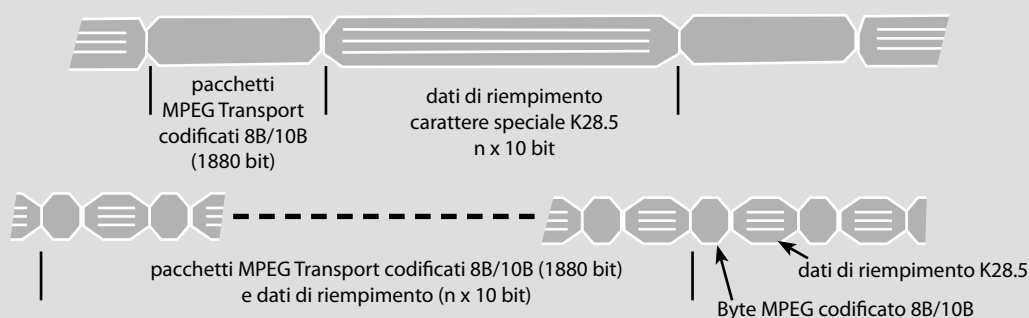



Fig. 9 - Esempio di applicazione dell'interfaccia DVB-ASI.
I pacchetti a lunghezza fissa che costituiscono il transport stream MPEG-2 comprendono 188 byte o 204 byte, nel caso siano protetti con il codice Reed Salomon. Il Pacchetto di Trasporto è costituito da una Header di 4 byte, il primo di sincronismo, e da un carico utile di 184 byte, per un totale di 188 byte: tale lunghezza fu scelta per ottenere una compatibilità con la struttura ATM (ciascuna cella ATM è costituita da 5 byte per la header, 1 byte di servizio e 47 byte di carico utile, l'equivalente in carico utile di 4 celle ATM, 4 x 47 byte, corrisponde quindi esattamente ai 184 byte di carico utile di un pacchetto del Transport Stream).

I byte sono codificati mediante un codice a gruppo che associa 10 bit ogni byte (8B/10B), in modo da consentire la rigenerazione del clock la rilevazione della presenza di errori; il sincronismo è ottenuto mediante una parola da 10 bit non compresa fra quelle generate dal codice.
I pacchetti di dati video possono essere forniti sia come burst continui di byte (stuttura in alto), oppure come byte isolati (struttura in basso).
A ciascun byte (nell'esempio in tabella sottostante 0100 0111) viene associato una di due possibili parole da 10 bit a seconda del valore del parametro RD (Running Disparity) che determina il rapporto tra il numero di zeri e uni durante la trasmissione. RD viene determinato in base al rapporto calcolato nel sub-block precedente.

carattere di informazione (8 bit)	0	1	0		0	0	1	1	1		
carattere di trasmissione (10 bit) con RD_+	1	1	1	0	0	0		0	1	0	1
carattere di trasmissione (10 bit) con RD_-	0	0	0	1	1	1		0	1	0	1

video non compresso, cioè quella supportata dalla SDI e di conseguenza dalla SDTI. E' quindi possibile sfruttare tale eccesso di capacità per consentire trasferimenti a velocità superiori al tempo reale. Ad esempio i videoregistratori che adottano il formato D10 comprimono il segnale video a 50 Mbit/s ed è possibile effettuare riversamenti, fra due videoregistratori, a velocità doppia rispetto al tempo reale: ad esempio un contributo di un'ora può essere riversato in mezz'ora.

Alcuni standard SMPTE definiscono la mappatura di un'ampia gamma di formati di videoregistrazione nel formato SDTI. Esistono inoltre ulteriori specifiche, non SMPTE, per consentire la mappatura di altri formati. Di fatto SDTI può essere utilizzata fra coppie di apparati operanti secondo i formati: DVcam, DVCpro25, DVCpro50, D9 (Digital-S) e *Transport Stream* MPEG-2, Betacam SX, e HDcam, D10 (IMX), DVCPRO HD, HD-D5.

3.2 Interfaccia asincrona DVB-ASI

E' l'interfaccia definita dal DVB [9] per consentire il trasferimento del *Transport Stream* MPEG-2 (figura 9) in modo elettricamente compatibile con la SDI (collegamenti unidirezionali, a 800 mV p-p, 270 Mbit/s, lunghezza del cavo, di qualità opportuna ed equalizzato, fino a 300 m).

ASI è l'acronimo di *Asynchronus Serial Interface*, esistono anche una interfaccia sincrona parallela (DVB-SPI) e una seriale (DVB-SSI) [10].

Si è visto che nell'interfaccia SDI esiste una relazione fissa tra la temporizzazione dei campioni video (27 MHz) ed il clock (270 MHz), quindi SDI è intrinsicamente una interfaccia sincrona.

Nel caso di SDTI e ASI, però, si trasporta un segnale compresso ed in questo caso può esserci ancora una relazione di sincronismo tra l'informazione video e il clock, ma spesso i sincronismi video sono indipendenti dal clock ricavabile dal flusso binario e sono invece ricavabili da opportune informazioni (*program clock reference*) contenute nel flusso di dati.

4. Interfaccia seriale per video ad alta definizione (HD-SDI)

Per l'alta definizione attualmente sono utilizzati soprattutto due formati d'immagine: HD-CIF previsto dalla Rac. BT.709 (1080 righe per quadro e 1920 pixel per riga) e quello previsto dalla Rac. BT.1543 (720 righe per quadro e 1280 pixel per riga). Il primo formato è oggi usato principalmente, in ambito televisivo, per la ripresa con scansione interlacciata, mentre la ripresa in formato progressivo trova applicazione soprattutto nell'ambito del cinema elettronico. Il secondo formato è previsto solo nella modalità di ripresa con scansione progressiva e per frequenze di immagine a 59,94 Hz, 30 Hz e 60 Hz, anche se sono stati sviluppati apparati operanti anche a 50 Hz.

In tabella 1 sono riassunti i bit-rate richiesti per le principali combinazioni di formato d'immagine e frequenza di ripetizione d'immagine.

Nel caso dei formati 1080i e 720p è sufficiente un bit-rate inferiore a 1,5 Gbit/s ed una apposita interfaccia

<table>
<tr><th>Formato</th><th>Scansione</th><th>Numero di righe attive per quadro</th><th>Numero di pixel per riga</th><th>Bit-rate [Gbit/s]</th></tr>
<tr><td>1080p60</td><td rowspan="4">progressivo</td><td rowspan="7">1080</td><td rowspan="7">1920</td><td>2,97</td></tr>
<tr><td>1080p59.94</td><td>2,97/1,001</td></tr>
<tr><td>1080p50</td><td>2,97</td></tr>
<tr><td>1080p25</td><td>1,485</td></tr>
<tr><td>1080i60</td><td rowspan="3">interlacciato</td><td>1,485</td></tr>
<tr><td>1080i59,94</td><td>1,485/1,001</td></tr>
<tr><td>1080i50</td><td>1,485</td></tr>
<tr><td>720p60</td><td rowspan="2">progressivo</td><td rowspan="2">720</td><td rowspan="2">1280</td><td>1,485</td></tr>
<tr><td>720p59,94</td><td>1,485/ 1,001</td></tr>
</table>

Tab. 1 - Bit-rate tipici per interfacce seriali HD

seriale specificata dalla norma SMPTE ST 292:2008 (*1.5 Gb/s Signal / Data Serial Interface*) definisce infatti una interfaccia seriale operante a 1,485 Mbit/s (per i formati a 60 Hz, 50 Hz, 30 Hz, 25 Hz e 24 Hz) e a 1,485/1,001 Mbit/s (per i formati a 59,94 Hz, 29,97 Hz e 23,98 Hz, frequenze di ripetizione d'immagine compatibili con il sistema NTSC).

Nei casi in cui la capacità di 1,5 Gbit/s non sia sufficiente, per supportare i formati che richiedono bit-rate più elevati (1080p, 4:2:2, 10 bit a 60 Hz, 59,94 Hz e 50 Hz oppure RGB, 4:4:4, a 12 bit), la norma ST 372:2009 (*Dual Link 1.5 Gb/s Digital Interface for 1920 x 1080 and 2048 x 1080 Picture Frames*) specifica un'interfaccia con una coppia di connessioni (*dual link*) consistente in due interfacce ST 292 in parallelo.

Poiché la soluzione con due connessioni pone problemi pratici rilevanti, è stata definita (ST 424:2006, *Television – 3 Gb/s Signal/Data Serial Interface*) una interfaccia a 2,97 Gbit/s che consente l'uso di un solo cavo (figura 10).

In tabella 2 sono indicate, a titolo d'esempio, le distanze raggiungibili, con cavo coassiale, mediante SDI e HD-SDI. Queste interfacce trovano applicazione nell'ambiente di produzione, per la televisione o per il cinema.

Tab. 2 - E' indicata la distanza di trasmissione calcolata per un tipo di cavo coassiale. Nell'esempio so riportati i dati del tipo 1694F della Belden. A riprova della ampia gamma di caratteristiche disponibili, si può osservare che lo stesso costruttore propone un tipo di cavo caratterizzato da una distanza che, ad esempio nel caso del bit-rate 270 Mbit/s, è limitata a 116 m ed un altro tipo che può arrivare a 554 m.

Applicazione	bit rate [Mbit/s]	Norme	m
NTSC composito	143	ST 259:2008 Annex A	457
PAL composito	177,3	ST 259:2008 Annex B	416
4:2:2 componenti a 4:3	270	ITU-R BT.656 ST 259:2008	326
4:2:2 componenti a 16:9	360	ITU-R BT.656 ST 259:2008	27
4:4:4 componenti a 16:9	540	ST 344:2000	229
HDTV	1485	ST 292:2008	87
HDTV scansione progressiva	2970	ST 424:2006	59

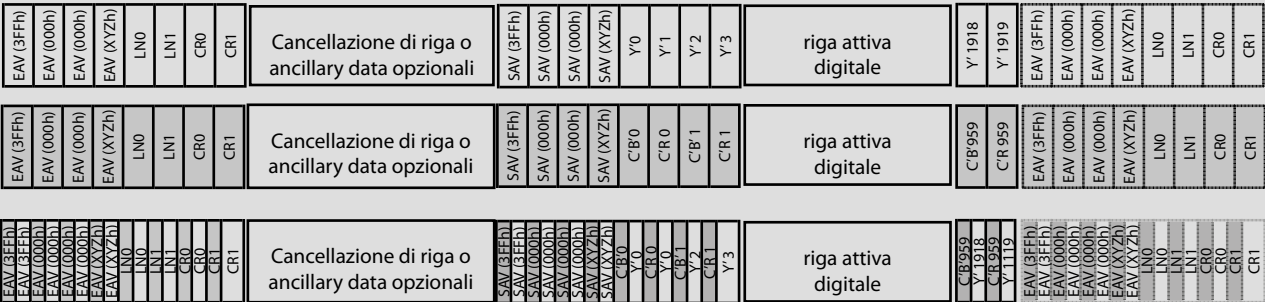



Fig. 10 - L'intefaccia a 3 Gbit/s (ST 424) prevede che il flusso seriale sia ottenuto a partire da due flussi, costituenti una interfaccia virtuale, di parole da 10 bit. Ciascuno dei due flussi è costituito da quattro aree: EAV e riferimento temporale, intervallo di cancellazione di riga, SAV e riferimento temporale, riga attiva (video o ancillary data). EAV e SAV sono costituite da quattro parole come indicato in figura 4, sono entrambe seguite da due parole che indicano il numero di riga e da due parole per rivelare eventuali errori (CRC).
Le parole di 10 bit costituenti i due flussi dell'interfaccia virtuale (in alto in figura) sono intercalate per formare il multiplex relativo all'interfaccia parallela riprodotto in basso in figura. I dati sono successivamente serializzati, con il bit meno significativo (LSB) di ciascuna parola trasmesso per primo. Il flusso complessivo ha un bit-rate pari a 2,97 Gbit/s (o 1,97/1,001 Gbit/s se la frequenza di quadro è quella NTSC). Lo schema di codifica di canale è NRZI con scrambler ed i polinomi generatori utilizzati sono gli stessi adottati per la SDI (figura 7). Nel caso di collegamento mediante cavo coassiale a 75 ohm e connettore BNC, il segnale ha un'ampiezza di 800 mV ±10%, dc offset 0,0 V ±0,5 V.

Fig. 11 - I connettori DVI possono essere di tipo DVI-I (supportano sia il segnale digitale che quello analogico) per singola connessione (single-link) o doppia connessione (dual-link). I DVI-D sono solamente per collegamenti digitali e i DVI-A solamente per collegamenti analogici.

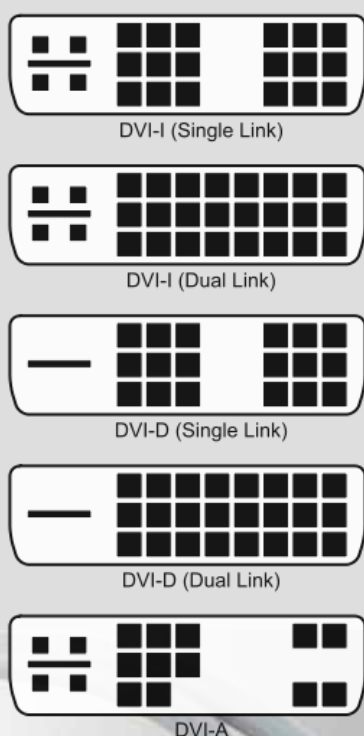



Fig. 12 - Cavo di interconnessione DVI (maschio) a sinistra e HDMI (maschio) a destra.

Altre norme, generalmente specificate da consorzi di produttori di apparati, si applicano agli apparati consumer: decoder, riproduttori, schermi e proiettori.

5. Dalla SCART alla HDMI

5.1 SCART

Nel 1977 fece la sua apparizione il connettore SCART, noto anche come *Peritel* e normalizzato per consentire l'interconnessione fra gli apparati audiovisivi (TV, VCR e, successivamente, DVD e console per i giochi) permettendo di trasferire i segnali video analogici RGB e compositi (dal 1980 anche S Video) e audio stereo.

5.2 DVI

La DVI (*Digital Visual Interface*) è un'interfaccia sviluppata nel 1999 [11] per consentire il trasporto di segnali video e la connessione anche in formato digitale di monitor a pannelli piatti e proiettori ai PC. Una connessione singola DVI consiste di quattro doppini (*twisted pair*) per il rosso, verde, blu e il clock e trasporta fino a 1,65 Mpixel/s, la lunghezzamassima del cavo è 5 m. Con una sola connessione a 60 Hz si può avere una risoluzione massima di 2098x1311 pixel e *aspect ratio* 10:9. Si può utilizzare una seconda connessione (sempre supportata dallo stesso connettore) per aumentare la risoluzione in numero di pixel oppure il numero di bit per pixel. Il connettore può ospitare anche i piedini per trasferire il video analogico, secondo lo standard VGA (figura 11).

5.3 HDMI

L'interfaccia HDMI (*High-Definition Multimedia Interface*) ha, per quanto riguarda la connessione digitali, caratteristiche simili alla DVI ma permette anche il trasporto del segnale audio. Sono reperibili cavi dotati sia del connettori DVI che di quello HDMI (figura 12). E' stata sviluppata da un consorzio di produttori (www.hdmi.org) specificatamente per il mercato consumer HDTV (figura 13).

La prima versione 1.0 è del 2002.

La versione 1.4 del maggio 2009 ha le seguenti caratteristiche: supporta almeno 7 formati 3D, supporta la scansione a 240Hz, supporta la comunicazione bidirezionale, che include la connessione Ethernet fino a 100 Mbps e lo streaming del flusso audio, funzionalità HEC, supporto al canale di ritorno per l'audio, supporto a risoluzioni fino a 4096 x 2160 pixel a 24 fps e 30Hz, funzionalità ACE, nuovo mini connettore a 19 pin.

La massima velocità di trasferimento video era nella versione 1.0 pari a 165 Mpixel/s, sufficiente per il formato HDTV 1080p o per UXGA (1600x1200), la versione 1.3 ha portato tale velocità fino al massimo di 340 Mpixel/s (per il tipo A, il doppio per il tipo B).

5.4 HDCP

Il termine HDCP (*High-bandwidth Digital Content Protection*) non ci si riferisce ad una interfaccia, bensì a un sistema di protezione, per evitare che pirati possano realizzare copie di contenuti in alta definizione.

Inizialmente fu sviluppato da Intel Corporation per evitare usi impropri dei contenuti transitanti sulle connessioni DVI e HDMI; le specifiche più recenti [12] permettono la sua applicazione per qualsiasi protocollo compatibile, incluso TCP/IP, e qualsiasi intefaccia via cavo o senza fili (*wired* o *wireless*), incluso Wi-Fi, Ethernet e USB.

Le specifiche HDCP sono proprietarie e i produttori di apparati devono richiedere una licenza alla organizzazione DCP per realizzare dispositivi che ne facciano uso. E' un sistema di gestione dei diritti DRM che applica la cifratura dei contenuti.

Un dispositivo HDCP, sia che sia la sorgente (*source*) sia la destinazione (*sink*) è dotato di una chiave privata HDCP. Le chiavi sono generate in modo che la coppia *source/sink* sia in grado di crittografare i dati che devono scambiarsi senza che la sorgente o la destinazione forniscano le rispettive chiavi.

Il meccanismo di protezione è basto su tre elementi.

- L'autenticazione dei ricevitori da parte del trasmettitore al momento in cui è richiesta la connessione.Grazie al protocollo di autenticazione il trasmettitore verifica che il il ricevitore ha la licenza per ricevere i contenuti.
- E' possibile che alcune chiavi private vengano ritenute non più valide e pertanto è realizzata una lista nera (*blacklist*) distribuita, ad esempio, attraverso apposite tracce registrate sui Blu-ray Disc, che permettono la revoca dei ricevitori che utilizzano tali chiavi, disabilitandoli dalla possibilità di riprodurre i contenuti in alta definizione.
- Le chiavi private, ciascuna di 56 bit, permette di criptare i contenuti che transitano sull'interfacce che collegano il trasmettitore con i ricevitori.

HDCP è una delle caratteristiche richieste, a partire dal gennaio 2005, dalla EICTA, ora DIGITALEUROPE, per poter identificare e commercializzare i display come HD ready e HD ready 1080p.

Pin	Segnale
Pin 1	TMDS Data2+
Pin 2	TMDS Data2 Shield
Pin 3	TMDS Data2–
Pin 4	TMDS Data1+
Pin 5	TMDS Data1 Shield
Pin 6	TMDS Data1–
Pin 7	TMDS Data0+
Pin 8	TMDS Data0 Shield
Pin 9	TMDS Data0–
Pin 10	TMDS Clock+
Pin 11	TMDS Clock Shield
Pin 12	TMDS Clock–
Pin 13	CEC
Pin 14	Reserved (HDMI 1.0-1.3c), HEC Data- (Optional, HDMI 1.4+ with Ethernet)
Pin 15	SCL (I²C Serial Clock for DDC)
Pin 16	SDA (I²C Serial Data Line for DDC)
Pin 17	DDC/CEC/HEC Ground
Pin 18	+5 V Power (max 50 mA)
Pin 19	Hot Plug Detect (All versions) and HEC Data+ (Optional, HDMI 1.4+ with Ethernet)

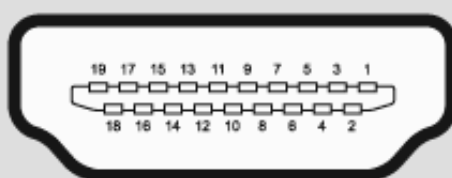



Fig. 13 - Connettore HDMI tipo A, da 19 contatti.
Quello di tipo B ha 29 contatti e permette il raddoppio della banda video (fino a 3840x2400 pixel. Il tipo C è per dispositivi portatili. I tipi D e E sono introdotti con la versione 1.4: il D mantiene i 19 contatti ma riduce le dimensioni, molto prossime a quelle di un mini USB, l'E è per l'Automtive Connection System.
Nella versione 1.3 e con connettore tipo A, il canale TMDS consente, per il video: velocità fino a 340 Mpixel/s, 4:4:4, RGB e YCBCR, (da 8 a 16 bit per componente) e 4:2:2 (fino a 12 bit per componente). Per l'audio: fino ad 8 canali, frequenza di campionamento: 32 kHz, 44,1 kHz, 48 kHz, 88,2 kHz, 96 kHz, 176,4 kHz, 192 kHz. Il canale DDC permette alla sorgente di interrogare il dispositivo di destinazione sulle proprie caratteristiche. Il canale CEC (Consumer Electronics Control), opzionale, è utilizzabile per funzioni di controllo remoto.

Già nel 2001, ancora prima dell'adozione commerciale del prodotto, fu messo in evidenza il fatto che il sistema di protezione era debole, il protocollo non era sufficientemente sicuro, e sarebbe stato possibile, mediante un processo di ingegnerizzazione inversa, identificare la *master key*. Tale chiave, una matrice di 40x40 parole da 56 bit, è quella utilizzata per creare le chiavi private distribuite da DCP per essere utilizzate dai dispositivi HDCP.

Nel settembre 2010 è stata effettivamente resa pubblica una chiave master e non è stato smentito il fatto che sia quella effettivamente utilizzata da DCP.

La conoscenza di tale chiave potrebbe consentire la realizzazione di dispositivi che neutralizzano la protezione: non è però banale realizzare dispositivi con tali caratteristiche e Intel ha dichiarato di voler perseguire, sulla base delle leggi che proteggono la proprietà intellettuale, chi realizzasse dispositivi realizzati sulla base di tale informazione.

E' quindi probabile che il sistema HDCP continui ad essere utilizzato dai produttori di contenuti, nonostante i dubbi sulla sua sicurezza.

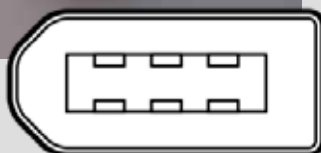
6. Interfacce per computer

Esistono numerose interfacce usate in ambito professionale per trasferire il video in ambito informatico. La più diffusa è denominata FireWire (da Apple), i.Link (da Sony) e IEEE 1394.

E' un'interfaccia bidirezionale utilizzata a partire dal 1995 e costituita da due doppini (*twisted pair*) ed un doppino opzionale per l'alimentazione. Il connettore è a 4 o 6 pin (figura 14).

Nell'applicazione più semplice supporta la comunicazione fra dispositivi a 100, 200 o 400 Mbit/s. La lunghezza massima del cavo è pari a 4,5 m. Possono essere interconnessi fino a 63 periferiche utlizzando una struttura di rete a hub. La specifica 1394b supporta fino a 900 Mbit/s e collegamenti in fibra ottica fino a 100 m e 3,2 Gbit/s.

SMPTE ST 396:2003 (*Television – Packet Format and Trans-mission Timing of DV-Based Data Streams over IEEE 1394*) ha definito il metodo per trasportare il video in formato DV, ma non è comunque garantita la compatibilità completa fra dispositivi che supportano lo standard IEEE 1394 perché esistono vari protocolli proprietari per trasportare audio, video e segnali di controllo.

Fig. 14 - Cavi IEEE 1394 con connettori a 6 e 4 pin.

Bibliografia

1. G.F. Barbieri: "Codifica Numerica del segnale video: standard per gli studi televisivi", Elettronica e Telecomunicazioni, n. 2, 1982, pp. 42-52
2. Recommendation ITU-R BT.601-6: "Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios" (1982-1986-1990-1992-1994-1995-2007).
3. Recommendation ITU-R BT.1361: "Worldwide unified colorimetry and related characteristics of future television and imaging systems" (1998).
4. Recommendation ITU-R BT.709-5: "Parameter values for the HDTV standards for production andinternational programme exchange" (1990-1994-1995-1998-2000-2002).
5. Recommendation ITU-R BT.656-5: "Interfaces for Digital Component Video Signals in 525-line and 625-line Television Systems operating at the 4:2:2 Level of Recommendation ITU-R BT.601", (1986-1992-1994-1995-1998-2007).
6. EBU - Recommendation R95: "Safe areas for 16:9 television production" (1999, 2000, 2008).
7. ISO/IEC 13818-2: "Information Technology - Generic coding of moving pictures and associated audio information: Video" (1996-2000).
8. Rec. ITU-R BT.1367 - Serial Digital Fiber Transmission Systems for Signals Conforming to Recommendations ITU-R BT.656, BT.799 e BT.1120).
9. ETSI TR 101 891 - Digital Video Broadcasting (DVB); Professional Interfaces: Guidelines for the implementation and usage of the DVB Asyncronuous Serial Interface (ASI), v1.1.1 (2001-02)
10. CENELEC EN 50083-9 - Cabled distribution systems for television, sonud and interactive multimedia signals. Part 9: Interfaces for CATV/SMATV headends and similar professional equipment for DVB/MPEG-2 transport streams.
11. Digital Visual Interface, rev. 1.0, 2 aprile 1999, www.ddwg.org.
12. High-bandwith Digital Content Protection System. Interface Independent Adaptation. Rev. 2.0, 23 Oct. 2008.

"Elettronica e Telecomunicazioni",
nata nel 1952 come "Elettronica e Televisione Italiana", è una rivista quadrimestrale di Rai Eri realizzata dal Centro Ricerche e Innovazione Tecnologica della Rai, sul cui sito è disponibile gratuitamente dal 2001.

Il Centro Ricerche e Innovazione Tecnologica (CRIT) della Rai nasce a Torino nel 1930 come "Laboratorio Ricerche" e dal 1960 ha sede in Corso Giambone 68. Successivamente assume la denominazione "Centro Ricerche" e, dall'ottobre 1999, quella attuale.
L'attività del Centro è coordinata dalla Direzione Strategie Tecnologiche.
Alla nascita, tra i suoi obiettivi ha la progettazione e realizzazione di impianti ed apparati di nuova concezione, non reperibili sul mercato. I profondi cambiamenti nello scenario delle telecomunicazioni hanno stimolato la trasformazione del Centro.
Ha ricevuto riconoscimenti a livello internazionale per i contribuiti forniti alle attività di studio e normalizzazione dei sistemi per la codifica dei segnali audio e video in forma digitale, allo sviluppo delle tecniche di compressione dei segnali attualmente alla base dei sistemi di trasmissione e registrazione dei segnali video, alla definizione degli standard di diffusione e trasmissione DVB.
Il Centro contribuisce all'evoluzione delle tecnologie relative al sistema radiotelevisivo e multimediale e supporta il Gruppo nelle scelte di indirizzo tecnologico e nella fase di sperimentazione e introduzione in esercizio di nuovi prodotti e sistemi. E' attivo in numerosi progetti finanziati in ambito europeo e nazionale e collabora con Università e Industrie per l'attività di ricerca, per la definizione dei nuovi standard e lo sviluppo dei nuovi servizi.

Rai Radiotelevisione S.p.A.
Centro Ricerche e Innovazione Tecnologica
Corso E. Giambone, 68 - I 10135 Torino
www.crit.rai.it